



Pro gradu -tutkielma
Fysiikan opettajan suuntautumisvaihtoehto

BEGREPPSKARTAN
SOM ALTERNATIV BEDÖMNINGSMETOD

Henrik Litonius

19.9.2013

Ohjaaja: Ismo Koponen

Tarkastajat: Heimo Saarikko, Ismo Koponen

HELSINGIN YLIOPISTO
FYSIKAN LAITOS

PL 64 (Gustaf Hällströmin katu 2) 00014 Helsingin yliopisto

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET – UNIVERSITY OF HELSINKI

Tiedekunta/Osasto – Fakultet/Sektion – Faculty/Section matematiske-Naturvetenskapliga/Fysik		Laitos – Institution – Department Fysik ämneslärautbildning	
Tekijä – Författare – Author Henrik Litonius			
Työn nimi – Arbetets titel – Title Begreppskartan som alternativ bedömningsmetod			
Oppiaine – Läroämne – Subject Fysik			
Työn laji – Arbetets art – Level Pro gradu -avhandling	Aika – Datum – Month and year 17.09.2013	Sivumäärä – Sidoantal – Number of pages 48	
Tiivistelmä – Referat – Abstract			
<p>Begreppskartor utvecklades ursprungligen på 1980-talet av Joseph Novak och Bob Gowin som ett sätt att strukturera och därmed få djupare förståelse för ny kunskap. Sedan dess har de utvecklats och undersökts som inlärnings-, utbildnings- och bedömningsmetod, inte minst av forskare som Maria Ruiz-Primo. Den här avhandlingen utgår till stor del från Ruiz-Primos utveckling av begreppskartan som ett verktyg för att bedöma elevers prestationer och undersöker huruvida traditionella prov kunde ersättas med begreppskartsuppgifter inom fysiken.</p> <p>Begreppskartor har en bevisligt positiv effekt vad gäller inläring med dem, men de används sällan i våra skolor och det är svårt att få elever att anamma en ny och ointuitiv inlärningsmetod. Att använda begreppskartan som en bedömningsmetod i klassen skulle knuffa eleverna mot ett mer begreppsbaserat tänkande och kunde hjälpa dem i sin förståelse av ämnet. Som provuppgift är begreppskartan snabb att göra och att rätta men ger läraren goda insikter i elevens förståelse för ett ämne. Ifall man kunde tänka sig att använda begreppskartan som provuppgift åtminstone diagnostiskt skulle det föra med sig en hel drös fördelar.</p> <p>Två klasser undersöktes, en i grundskolan på årskurs sju, och en kurs i gymnasiet. Bägge grupper fick lära sig grunderna i att göra begreppskartor på förhand, gymnasisterna hade lite erfarenhet av det också från tidigare. Båda grupperna fick sedan under en kurs skriva ett traditionellt prov samt göra en begreppskarta av materialet de nyligen gått igenom. Begreppskartorna bedömdes med en femstegs-modell utvecklad av Ruiz-Primo och Shavelson och gavs ett vitsord beroende på hur väl de jämfördes med en expertkarta. Resultaten undersöktes skilt för de två grupperna och granskades för svårighetsgrad, korrelation och överensstämmelse enligt en statistisk metod utvecklad av Bland och Altman 1986.</p> <p>Svårighetsgraden på uppgiften fanns vara lämplig, bägge gruppernas medeltal var något lägre för begreppskartan än för det traditionella provet vilket kan motiveras med att eleverna trots allt har större erfarenhet av traditionella prov än begreppskartor. Korrelationen för grundskolegruppen fanns vara god, 0,825 medan den för gymnasiet var betydligt svagare, 0,412. Bland-Altman metoden gav vidare negativa resultat för gymnasiet med mycket stora kast mellan de enskilda elevernas prestationer i de två uppgifterna. Grundskolegruppen presterade lite mer konsekvent men visade en trend där de verkligt svaga och de verkligt starka eleverna gynnades av begreppskartsuppgiften medan eleverna med medeltal kring 7 gjorde sämre ifrån sig än i det traditionella provet.</p> <p>Korrelationen för grundskolan är så pass stark att det är tänkbart att begreppskartan kunde användas som en bedömningsmetod inom grundskolan. Grundskolans begränsade matematik gör att också en stor del av naturvetenskaperna är fenomen- och begreppsbaserade snarare än baserade på problemlösning. I gymnasiet är det tvärtom, största delen av gymnasiekursen inom fysiken går ut på matematisk uträkning av fenomen, inte på att kunna förklara dem med ord och förstå samband. Som en följd av detta är begreppskartorna mer användbara som en alternativ bedömningsmetod i grundskolan än i gymnasiet.</p>			
Avainsanat – Nyckelord – Keywords Begreppskarta, bedömningsmetod, diagnostik			
Säilytyspaikka – Förvaringställe – Where deposited Helsingfors Universitet			
Muita tietoja – Övriga uppgifter – Additional information			

Innehållsförteckning

1 Inledning.....	5
2 Teoretisk bakgrund	6
2.1 Novak och Gowin lägger grunden för begreppskartan.....	6
2.2 Senare teoretisk utveckling och kritik.....	7
2.3 Hur mäts begreppskartor?.....	9
2.4 Pålitligheten och relevansen i begreppskartor	9
2.4.1 Uppgiften.....	10
2.4.2 Svarsmediet.....	12
2.4.3 Bedömningen	13
2.4.4 Tre exempel på bedömningskriterier av begreppskartor	15
3 Studiens syfte, begreppskartan i bruk som bedömningsmetod	18
4 Material.....	20
4.1 Experimentets uppbyggnad.....	20
4.2 Begreppsval och expertkarta	21
5 Analys	23
5.1 Kartuppgiften som helhet.....	24
5.2 Kvalitativ analys	24
5.2.1 Begreppskartorna från grundskolegruppen.....	27
5.2.2 Begreppskartorna i gymnasiet	31
5.3 Kvantitativ analys, en jämförelse mellan begreppskartan och traditionella prov som mätmetod.....	34
5.3.1 Svårighetsgrad och medeltal	35
5.3.2 Korrelation	35
5.3.3 Överensstämmelse med de traditionella förhören	39
6 Diskussion	41
Källförteckning	45
Bilaga 1: Lista på begrepp samt förhört för grundskolegruppen.....	46
Bilaga 2: Lista på begrepp samt förhört för gymnasiegruppen.....	48
Bilaga 3: Provresultat	50

Bildförteckning

Bild 1 Ett exempel på en begreppskarta.	6
Bild 2 Ett exempel på en cyklisk begreppskarta.	8
Bild 3 Hierarkisk begreppskarta poängsatt enligt principerna framlagda av Novak och Gowin	15
Bild 4 Elev P från grundskolan. Eleven har inte anammat begreppskartans struktur eller så är elevens kunskapsbas mycket lösryckt och splittrad.	28
Bild 5 Elev C från grundskolan. Begreppskartan har en stark och relativt välbalanserad nätverksstruktur där majoriteten av länkarna är mycket välformulerade.	29
Bild 6 Elev K från grundskolan. Typisk kedjestruktur där de flesta begreppen har en eller två länkar kopplade till sig.	30
Bild 7 Elev D från grundskolan. Kartans struktur är i sig bra men länkarna är svaga och innehåller många missförstånd.	31
Bild 8 Elev K från gymnasiet. Kartan i sig är mycket starkt centrerad runt "Resistans" som begrepp medan länkarna visar på en förståelse som främst grundar sig på matematisk problemlösning och formler.	33
Bild 9 Elev B från gymnasiet. Kartan är svag både till innehåll och struktur. Få noder är kopplade till mer än två andra noder, vissa svävar fullständigt fritt och de felaktiga länkarnas antal är högt.	34

Grafförteckning

Graf 1 Resultat av provet och begreppskartan för grundskolan.	37
Graf 2 Resultat för provet och begreppskartan för gymnasiet	37
Graf 3 Bland-Altman-graf för grundskolan	40
Graf 4 Bland-Altman-graf för gymnasiet.	41

1 Inledning

Skolvärlden i allmänhet och lärare i synnerhet kämpar ständigt med bristande tidsresurser. Kursinnehåll har en tendens att öka samtidigt som kraven på mångsidig, individuell och spännande undervisning växer medan timresurserna hålls lika eller krymper. Förutom att presentera lärostoffet måste tid reserveras för repetition, övningar och inte minst bedömning av de enskilda eleverna. Utvärderingsövningar i form av kursprov och förhör är ett ypperligt sätt att testa elevernas kunskap och ger eleverna en stark motivation att läsa och lära sig stoffet, men provtillfällena är ofta bortkastade ur inlärningssynpunkt: endast sällan kan man säga att eleven lär sig något nytt under provskrivningen. Likaså har stora delar av proven en tendens att bygga på *utantillinlärning* så läraren får inte heller några större insikter i elevens *förståelse*. Provrättningen och genomgång av de samma kostar läraren tid som annars kunde användas till vettigare förberedelser och ytterligare timresurser. Ifall ett bedömningsverktyg kunde framställas som är snabbt, pålitligt och samtidigt kan ge både läraren och eleven nya insikter skulle det vara en värdefull resurssparare. Den här studien har som syfte att undersöka huruvida begreppskartor, en pedagogisk metod utvecklad av Novak och Gowin på 1980-talet (Novak & Gowin 1984), kunde vara en sådan bedömningsmetod och hur väl den lämpar sig för grundskolan och gymnasiet.

I forskningslitteraturen har mycken kritik framförts gentemot de slags prov och mätmetoder som vanligtvis används av skolsystem världen över, inte minst av Maria Ruiz-Primo och Richard Shavelson. De menar att kunskap består av tre nivåer, påståenden, processer och strategier. Dessa kan man förklara som att veta att något är på ett visst sätt (påstående), att veta hur man gör något (process) samt att veta när kunskapen skall tillämpas (strategi) (Ruiz-Primo, 2000). Ett exempel ur fysiken kunde vara kunskapen om acceleration på grund av gravitation. Påståendet skulle vara att veta att alla föremål på jordens yta dras mot jordens centrum med en acceleration på ca $9,81 \text{ m/s}^2$. Processuell kunskap skulle vara att kunna räkna ut hastigheten för till exempel ett fritt fallande objekt. Strategisk kunskap skulle vara att förstå att gravitationen spelar en stor roll i frågan om kaströrelser men inte just kommer att påverka en bil som rör sig rakt framåt på en vågrät väg. Vanligtvis brukar prov endast mäta en av dessa tre kunskapsformer, nämligen påståendena (Ruiz-Primo, 2000).

I det här skedet måste det påpekas att den skolvärld Ruiz-Primo verkar vara bekant med och har undersökt är till synes mycket mer benägen att använda sig av flervalsprov än av essäuppgifter och räkneuppgifter som är vanligare i det finska

skolsystemet. Det är ändå önskvärt att skapa ett system där inte bara utantillinlärning utan också förståelse av ett ämne genom kopplingar mellan de olika begreppen, även för oss. Essäsvår fyller delvis dessa kriterier men att rätta essäer är tidskrävande och skrivandet kan vara onödigt krävande om det endast är kopplingar och helheter man är ute efter. I sådana fall presenterar begreppskartor en möjlighet för läraren att se hur en elev förstår ett givet ämne utan att behöva plöja igenom flera sidor material. Rätt uppbyggt är begreppskartan även lättare att korrigera objektivt då till exempel elevens språkkunskaper spelar en mindre roll i dennes förmåga att presentera sitt material med det lilla aber att elever med mycket svagt språk eller med en språkbakgrund som inte stöder begreppskartans uppbyggnadssystem kan vara onödigt handikappande då de tvingas skapa begreppskartor (İnceç, 2009).

Räkneuppgifter har en tendens att dominera speciellt gymnasiefysiken och ger så tillvida en bra insikt i elevens processuella kunskap. Att klara av att räkna matematiska problem, speciellt sådana som övats tidigare, är dock ingen garanti på att eleven i fråga har verklig förståelse för fenomenet, endast att hen kan bolla med formlerna. På grundskolenivå har elevernas matematiska förmåga å andra sidan en tendens att vara så pass svag att de flesta matematiska problem kan vara onödigt krävande. På bägge nivåer är alltså rent matematiska uppgifter och problemlösning inte garanterade att ge läraren en bra bild av hur eleven greppar ämnet i sin helhet. Också här har begreppskartan något att ge.

Pedagogik är en verktygslåda för läraren. Något mirakelverktyg som skulle passa för alla situationer, individer och ämnen finns inte, istället finns det en hel drös metoder som alla har sina styrkor och svagheter. Förhoppningsvis kunde begreppskartan som bedömningsmetod läggas till som ett verktyg.

2 Teoretisk bakgrund

2.1 Novak och Gowin lägger grunden för begreppskartan

Begreppskartor som koncept presenterades första gången av Joseph Novak och Bob Gowin i deras verk *Learning how to learn* (Novak & Gowin 1984). I grund och botten är begreppskartor ett sätt att grafiskt framställa kunskap på ett sätt som visar förhållandena mellan de olika delarna av ämnet i fråga. De byggs upp av ett antal påståenden där begrepp eller koncept knyts samman med länkande pilar, se Bild 1 för ett exempel (Ruiz-Primo & Shavelson, 1996).

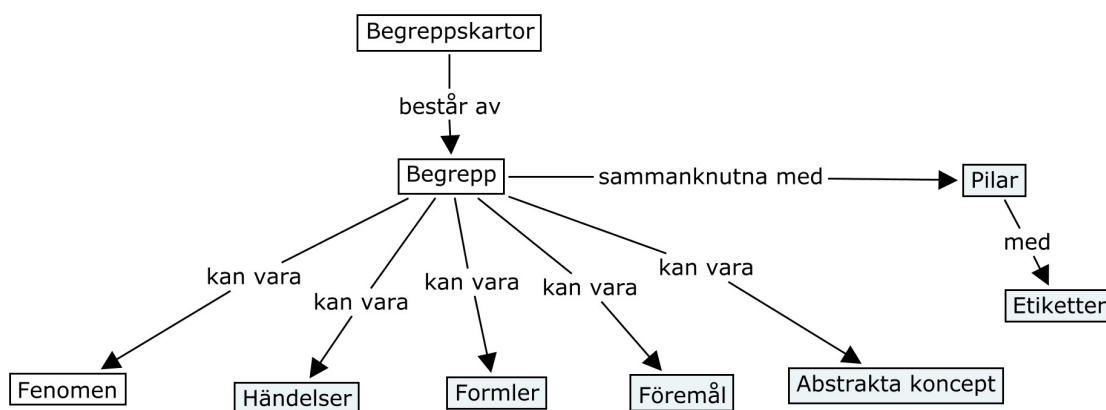


Bild 1 Ett exempel på en begreppskarta.

Metoden utvecklades efter att Novak och Gowin under sina universitetsföreläsningar märkte att de metoder och system för förståelse och inläring de som experter på området använde sig av började användas av deras studenter. Som följd av detta började de aktivt lära ut dessa system och testa dem, först på universitetsstuderanden under 1970-talet och sedan under 1980-talet på yngre elever i låg- och högstadiet (Novak & Gowin 1984, s. 9). Tanken med begreppskartan var att den skulle hjälpa både elever och lärare att få en djupare förståelse för innehållet i undervisningsmaterialet (1984 s. 1).

Novak och Gowin baserar teorin bakom begreppskartorna på David Ausubels konstruktivistiska kognitiva teorier från 1960-talet. Enligt den kan riktig inläring, alltså *meningsfull* inläring, endast ske då ny kunskap kopplas till och läggs på gammal. Det här är inte något man kan tvinga någon till utan det är en aktiv process där eleven själv bygger upp sin förståelse för världen runt om sig. Motsatsen till sådan väluppbyggd och meningsfull kunskap är enligt Ausubel utantillinläring, där eleven helt enkelt lär sig det stoff som presenteras utan att göra några kopplingar till sina egna erfarenheter (1984 s. 7). Novak och Gowin konstaterar: *"kunskap inte kan upptäckas som guld eller olja utan måste byggas som bilar eller pyramider."* (1984 s. 4).

2.2 Senare teoretisk utveckling och kritik

Begreppskartan som Novak och Gowin presenterade den har kritiserats en hel del med åren. Bland annat ansåg de att begreppskartorna måste vara hierarkiskt uppbyggda, så att allmännare begrepp är högst upp på en karta och mer specifika begrepp finns längre ner. Den här synen bygger på Ausubels kunskapsteori som kräver att all kunskap skall vara ordnad så. Kritiker har hävdat att man likväl kan härleda begreppskartan ur Deeses teori om minne genom association (Ruiz-Primo & Shavelson, 1996). Både Ausubels och Deeses teorier går ut på att meningsfull inläring sker genom att inordna nya fakta med gamla, men Novak och Gowin accepterar inte icke-hierarkiska metoder att bygga upp begreppskartor på. De

menar att till exempel cykliska begreppskartor inte stämmer överens med deras utvalda kunskapsteori (se Bild 2) (1984 ss. 37-39).

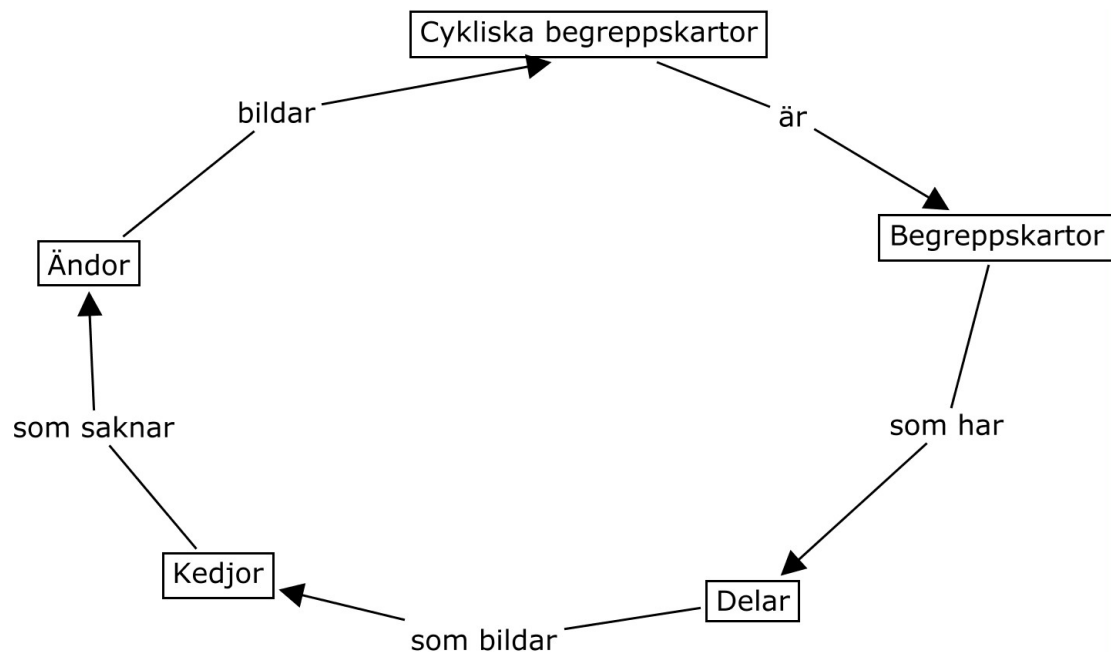


Bild 2 Ett exempel på en cyklisk begreppskarta.

Detta har ifrågasatts på senare tider och starka argument har lagts fram för användningen av cykliska begreppskartor för att presentera dynamiska, snarare än statiska processer, som till exempel vattnets kretslopp, fotosyntes eller ekonomiska kretslopp (Safayeni, Derbentseva, & Cañas, 2005). Då den cykliska begreppskartan utnyttjas så är det dock främst som ett verktyg för inläring och förståelse medan Novak och Gowin i sitt ursprungsarbete presenterar begreppskartan främst som ett diagnostiskt verktyg för att ge läraren en inblick i den enskilda elevens förståelse av det presenterade materialet (1984 s. 23), vilket kan vara orsaken till att de ställer sig negativt till att presentera materialet cykliskt.

Det kan ses som en aning motsägelsefullt att Novak och Gowin å ena sidan menar att styrkan hos begreppskartan ligger i den enskilda elevens förmåga att konstruera sina egna kartor för att kartlägga sin förståelse av ett ämne men samtidigt vill ordna all kunskap i hierarkier enligt Ausubels kunskapsteori. Att tvinga elever att forma sina tankar i hierarkiska system är mycket problematiskt. Ett sådant statiskt, hierarkiskt tänkesätt är inte lämpat för att presentera alla tänkbara förhållanden på ett naturligt sätt (Ruiz-Primo & Shavelson, 1996). Novak och Gowin själva erkände att den valda hierarkin mycket väl kan vara arbiträr (1984 ss. 16-17). Novak & Gowin talar också om "*gummikartor*" (1984 s. 17) där samma ämne kan presenteras med flera av begreppen som högsta hierarkiska nivå beroende på smak och tycke. Att då stenhårt hålla fast vid att alla ämnen ändå måste presenteras hierarkiskt är inte helt lätt att motivera. Frågan är om inte deras starka tilltro till Ausubels hierarkiska kunskapsstruktur i det här fallet kan ha inverkat negativt och

om det inte varit bättre att från första början tillåta mer naturliga icke-hierarkiska strukturer. Som det nu presenteras är hela deras metod att till exempel bedöma begreppskartor grundad på att de är uppbyggda med klara nivåer. Sedan Novak och Gowin lade fram sin metod har ett antal forskare utnyttjat och experimenterat med den, sammanlagt har 128 olika kombinationer av uppbyggnad, svarsform bedömning med mera använts (Nicoll, 2001), några av dem presenteras nedan.

2.3 Hur mäts begreppskartor?

Begreppskartor används främst på fyra olika sätt. Som ett verktyg för undervisning, som ett hjälpmedel för inläring, som en metod för kursplanering samt som ett instrument för att mäta elevers kunskapsnivå och förståelse (McClure, Sonak, & Suen, 1999). Av dessa olika användningsområden kommer denna undersökning att fokusera på det sistnämnda, begreppskartan som en mätmetod, samt hur väl den korrelerar med elevers allmänna prestationer och resultat.

Ifall man använder begreppskartor för sin egen inlärnings skull eller för att gestalta sina idéer så att de skall vara enkelt tillgängliga senare är det inte så noga vilka tekniker eller system man använder, huvudsaken är att informationen byggs upp på ett sätt som hjälper en själv att minnas den senare. Men då begreppskartor skall användas för att mäta kunskap är det viktigt att formulera klara regelverk för att försäkra sig om att mätningen inte blir allt för subjektiv. Ruiz-Primo och Shavelson har formulerat problematiken så att för att begreppskartor skall kunna anses användbara måste de vara både pålitliga samt relevanta (Ruiz-Primo & Shavelson, 1996). Med pålitlighet avses förmågan av olika bedömare att komma till liknande resultat och med relevans avses att uppgiften måste vara sådan att man kan anta att den på något vis är kopplad till elevens kunskap. Olika metoder att bygga upp begreppskartsuppgifter har olika nivåer av pålitlighet och relevans.

Under de närmare tre årtionden man forskat i begreppskartor har ett otal olika former av mätmetoder presenterats. Ruiz-Primo och Shavelson har systematiserat dessa metoder i ett antal studier under 1990-talet. Deras syn är att varje uppgift vars syfte är att producera en begreppskarta består av tre delar, uppgiften, svarsformen samt bedömningen (Ruiz-Primo & Shavelson, 1997). Saknas någon av dessa är det inte frågan om ett sätt för läraren att mäta elevens kunskaper.

2.4 Pålitligheten och relevansen i begreppskartor

Sedan de utvecklades i början av 1980-talet har begreppskartor använts för ett antal olika syften, från planering av kurser till inläringsteknik. En storskalig metastudie visade att användning av begreppskartor för inläring var effektivare än passivare studietekniker som lyssnande, läsande etc. Det fanns till och med vissa svaga indicier

att begreppskartor skulle vara effektivare än till exempel sammandrag, men inte tillräckligt för att kunna dra några slutsatser (Nesbit & Adesope, 2006). Studier som Nesbits har gjorts om inläring genom begreppskartor och ett otal studier har gjorts angående hur konsekvent och trovärdigt en lärare kan bedöma begreppskartor (Ruiz-Primo, Shavelson, McClure för att nämna några forskare på området) men mycket få studier har tagit upp vad begreppskartor mäter och hur relevant det är.

Ifall en lärare väljer att använda begreppskartan som en mätmetod och genom den försöka bedöma kunskapsnivån hos enskilda elever är det viktigt att veta huruvida begreppskartan överhuvudtaget kan antas mäta något relevant. McClure et al. nämnde tre fallgropar och problem att ta i beaktande ifall begreppskartan används för att bedöma elever. För det första kan eleven ha bristfälliga kunskaper i att skapa begreppskartor. Ifall du inte förstår principen kommer du inte att kunna uttrycka dig ordentligt genom begreppskartan och bedömningen av din skapelse kommer inte att vara ett verkligt mått på din kunskap i ämnet. För det andra kan det finnas skillnader i bedömarnas kunskapsnivå som leder till att olika bedömare anser olika påståenden olika riktiga och relevanta. Slutligen kan bedömningsmetoden öppna för inkonsekvens vilket leder till att också en enskild bedömare kan ge olika mängder poäng för samma påstående (McClure et al., 1999). Dessa tre fallgropar och hur jag har försökt undvika dem i den här studien diskuteras närmare i kapitlet om undersökningens uppbyggnad.

2.4.1 Uppgiften

Uppgiften eleven får framför sig hänvisar främst till den frihetsgrad denne har då begreppskartan skall byggas upp (Ruiz-Primo & Shavelson, 1997). Dessa kan gå från mycket kontrollerade svarsformer som "fyll-i-luckan"-uppgifter där ord skall tas från en lista och sättas in antingen som begrepp eller länkar, till mycket fria uppgifter där eleven endast presenteras med ett tema och sedan själv får bygga upp sin begreppskarta. För att nå en förståelse för problematiken i uppgiftsgivandet skall vi granska några exempel på olika frihetsgrader.

Den allra bästa insikten i en elevs psyke kan man antas få ifall eleven ges fria händer i fråga om svar, ifall hon får bygga sin begreppskarta hur stor eller liten som helst, i vilken form som helst och användande sig av vilka som helst begrepp och länkar för att bygga upp sina påståenden. Denna frihet kommer dock med ett högt pris i form av svårigheter i bedömningen. Fria kartor kan vara mycket svåra att bedöma rättvist ifall de skiljer sig anmärkningsvärt mycket från expertkartorna. Tanken är ändå att eleven med tiden får en kunskapskonstruktion som allt mer börjar likna expertens. Det finns dock teorier som menar att kravet kan vara för

kan bli överväldigade av kravet att producera en helgjuten bild av ett kunskapsområde vilket kan leda till att deras kartor ger en onödigt dålig bild av deras kunskapsnivå (Ruiz-Primo, 2000).

Motsatsen till den fria kartan är den där eleven endast får välja vilka ord från en lista som skall användas i vilka luckor på en begreppskarta. Dessa är mycket lätta att bedöma, i många fall kan bedömningen skötas av en dator (Yin, Vanides, Ruiz-Primo, Ayala, & Shavelson, 2005), men har visat sig lämna mycket att önska i form av vad de egentligen berättar om elevens kunskapsnivå. Ruiz-Primo och Shavelson kom till slutsatsen att de egentligen inte kan anses öppna något relevant fönster in i elevens sinne (Ruiz-Primo & Shavelson, 1996). Dyliga uppgifter har en tendens att ge nästan fulla poäng åt en mycket stor del av eleverna eftersom de av uppgiften leds mot det riktiga svaret. Vare sig det är begreppen, länkarna eller en blandning av bägge som skall väljas från en lista kan det ifrågasättas ifall dessa ger någon meningsfull information om elevernas kunskapsstrukturer (Yin et al., 2005).

Forskningen i begreppskartor har undersökt skillnaden mellan uppgifter där begreppen är givna men de länkande fraserna måste skapas på egen hand (Created linking phrases), eller C-kartuppgifter samt uppgifter där både begreppen och de länkande fraserna är givna (Selected linking phrases), eller S-kartuppgifter (Yin et al., 2005). Åter visade det sig att den friare uppgiften var bättre på att skilja mellan elevernas förståelse medan den mer kontrollerade var enklare att korrigera vilket gjorde den mer lämpad för storskaliga undersökningar. Det visade sig också, att den större valfriheten i C-kartorna gjorde att eleverna kunde utföra uppgifterna snabbare. Eleverna i den aktuella studien fick 9 begrepp att länka samman. I medeltal tog det eleverna, som måste skapa sina egna länkar, mindre än 10 minuter att göra uppgiften medan de som konstruerade S-kartor tog närmare en kvart timme på sig. Orsaken torde vara den, att då eleven valde två begrepp tänkte hen automatiskt på deras inre förhållande. I C-kartornas fall kunde eleven sedan skriva ner detta förhållande medan i S-kartornas måste denne ännu gå igenom de givna länkarna för att se vilken av dem som bäst stämde överens med den egna uppfattningen (Yin et al., 2005). Det verkar alltså som om det enda tillfället då det är värt att använda S-kartor är vid stora, hela årskurser omfattande undersökningar. För klassrumsbehov är det inte värt att ge eleverna de länkande fraserna.

Undersökningar har alltså visat att det av praktiska skäl är värt att ge eleverna begränsningar i form av vilka begrepp de skall använda (Ruiz-Primo & Shavelson, 1997). Senare studier har också visat att det inte lönar sig att påtvinga eleverna begreppskartans struktur i form av en fyll-i-kartan-metod eller ens att tvinga fram en hierarkisk eller icke-hierarkisk karta (Ruiz-Primo & Shavelson,

1996). Märk att detta går rakt emot den ursprungliga tanken bakom begreppskartor som den presenterades av Novak och Gowin. Deras syn, starkt baserad på Ausubels teori om strukturerad inläring kräver hierarkiska begreppskartor (Novak & Gowin, 1984). Senare forskning har snarare baserat begreppskartorna på Deeses teori om associativ inläring (Ruiz-Primo & Shavelson, 1996). Kvarstår alltså frågan om hur de givna begreppen skall väljas ut. Den frågan undersöktes av Ruiz-Primo och Shavelson 1996.

I en forskning publicerad 1997 undersökte Ruiz-Primo och Shavelson resultatet av två olika metoder, en där eleverna gavs 10 begrepp och fick i uppgift att bygga en karta av dessa och en annan där eleverna själva fick välja 10 begrepp (i praktiken 7, eftersom 3 begrepp var givna i och med området) och sedan göra en karta baserad på dem (Ruiz-Primo & Shavelson, 1997). Resultaten var nästan lika för de bägge grupperna, men då eleverna hade valt begreppen själva valde de ofta mer eller mindre irrelevanta eller allt för vaga begrepp att arbeta med. Detta gjorde att de som rättade måste dela ut poäng också för sådana länkar som inte i sig visade på någon djup kunskap. Slutsatsen var att det blir allt för stor variation på kartorna och allt för luddiga resultat ifall eleverna själv får välja vad de sätter in (Ruiz-Primo & Shavelson, 1997).

Baserat på de ovannämnda forskningsresultaten verkar alltså den ideala balansen mellan utförbara och mätbara uppgifter som samtidigt berättar något relevant om elevens kunskapsnivå ligga i att begränsa elevens val av begrepp men ge hen fria händer i frågan om länkar och strukturer. Det här kommer i sin tur att påverka eventuell bedömning av begreppskartorna, vilket ännu skall granskas efter några ord om svarsmedlet.

2.4.2 Svarsmediet

Av de tre delarna av en begreppskartsbedömning, uppgiftens uppbyggnad, svarsmedlet och bedömning kan svarsmedlet te sig minst viktigt för resultatet och vars följder inte vederligen har studerats grundligt. Det är frågan om i vilket format eleven producerar sin begreppskarta. Oftast är det genom penna och papper, men det kan också vara med hjälp av en dator eller något annat medium, som magneter på krittavlan eller post-it-lappar på väggen. I en undersökning lät Gayle Nicoll forskningsassistenterna göra upp begreppskartor baserat på essäer eller intervjuer med eleverna (Nicoll, 2001). Denna sistnämnda metod kan måhända ge mycket goda och exakta resultat. Nicolls undersökning kunde ta i beaktande elevens mån av säkerhet för diverse påståenden. Se kapitel 2.4.3 för en närmare beskrivning av Nicolls et. al. experiment. Den formen av mycket personlig interaktion mellan

mycket begränsade fall som specialundervisning eller dylikt. Magneter och klisterlappar har fördelen att de är storskaliga och lätta att handgripligen flytta omkring men skapar inte något permanent och kan därför vara svåra att bedöma och hänvisa tillbaka till. Champagne et al. använde sig 1978 av kort med olika begrepp skrivna på dem. Dessa skulle först sorteras i kända och okända begrepp och sedan skulle eleverna försöka ordna dem enligt hur de förhöll sig till varandra (Ruiz-Primo & Shavelson, 1996), en metod som på grund av sin handgriplighet kan vara ytterst användbar speciellt för yngre elever.

De vanligaste metoderna för att utföra begreppskartsbedömningar verkar vara genom att använda datorer eller penna och papper. Någon stor och tydlig skillnad mellan dessa två metoder har inte funnits. Den största fördelen med att göra en karta på dator torde vara enkelheten i att göra förändringar samt möjliggörandet av datoriserade bedömningar och analyser. Då datorer har blivit en allt vanligare syn också i skolor och då billiga eller kostnadsfria program för skapandet av begreppskartor som CMAPtools blivit allt tillgängligare torde det även bli vanligare att göra begreppskartor på ett datoriserat medium medan de handgjorda kartorna kan användas för snabba små kartor främst för eget bruk.

2.4.3 Bedömningen

Minst lika viktigt som hur begreppskartan byggs upp är hur läraren bedömer den. Trots att bedömningen delvis kommer att bestämmas av vilka begränsningar som lagts på uppgiften samt av den form svaret getts i är det till stor del helt upp till bedömaren att ställa upp sina egna kriterier. Ruiz-Primo och Shavelson presenterar bedömningskriterierna för ett 20-tal olika forskningar (Ruiz-Primo & Shavelson, 1996). McClure et al. undersökte sex slags bedömning samt jämförde dem med varandra (McClure et al., 1999) liksom också İnceç tio år senare (İnceç, 2009), så det finns bedömningskriterier att välja emellan. Grovt sett kan man dela in dessa kriterier enligt de begränsningar de kräver av uppgiften samt hur mycket frihet de ger bedömaren. Dessa sträcker sig från mycket lösa kriterier, som en holistisk bedömning av vad eleven visat sig förstå, till mycket strikta, som Novak och Gowins ursprungliga bedömningskriterier. Både McClure och İnceç var speciellt intresserade av hur mycket olika bedömares åsikter skiljde sig från varandra och alltså hur pålitliga de olika systemen kunde anses vara.

Överlag kan bedömningskriterierna delas in i tre grupper. Den vagaste av dessa är den holistiska, där bedömaren försöker skapa sig en bild av elevens kunskapsstruktur som helhet och ge denna ett poängtal från 1 till 10 (İnceç, 2009), (McClure et al., 1999). Denna metod är föga förvånande mycket snabb, i medeltal

de mest tidskrävande metoderna men har också en hög tendens till godtycke inte bara mellan olika bedömare utan även för enskilda bedömare mellan olika kartor (McClure et al., 1999).

Med jämförbar exakthet kan begreppskartor bedömas antingen enligt relationella eller strukturella kriterier. Relationella kriterier betyder att de olika påståendena, två begrepp sammankopplade med en beskrivande länk, bedöms separat enligt givna kriterier och summan av alla dessa enskilda poängsättningar ger begreppskartans slutgiltiga poängtal (İngeç, 2009). Två sådana metoder kommer att granskas nedan, en utvecklad av Nicoll et al. (Nicoll, 2001) samt en utvecklad av Ruiz-Primo och Shavelson (Ruiz-Primo & Shavelson, 1997). Strukturell bedömning innebär förutom en analys av de enskilda påståendena även att strukturen på begreppskartan spelar en roll. Det vanligaste exemplet på detta är det ursprungliga bedömningssättet som det framfördes av Novak och Gowin (Novak & Gowin, 1984). Eftersom de strukturella bedömningskriterierna är beroende av att eleven producerar en viss slags karta måste dessa uppgifter oftast vara mer begränsade än de relationella. Överlag kan man säga att det finns en slags Heisenbergs osäkerhetsprincip över bedömningen av begreppskartor. Ju exaktare bedömning man kan åstadkomma, desto mer måste uppgiften begränsas och desto mindre berättar kartan om vad eleven verkligen förstår. Friare uppgifter ger å andra sidan goda insikter i elevens kunskaper men är förrädiskt svåra att bedöma objektivt och tidseffektivt.

Då de enskilda kartorna bedömts kvarstår att analysera resultaten. Vanligast är att använda sig av en expertkarta, alltså en karta som är gjord av en mer kunnig person, oftast läraren, enligt samma kriterier som eleverna hade. Den här kartan anses vara det eleverna skall sträva efter att uppnå. Expertkartan bedöms sedan enligt samma kriterier som de övriga kartorna och ger bedömarens en uppfattning om hur långt eleven har kommit, oftast i form av ett procenttal. Märk väl att det är fullt möjligt att en elev får fler poäng än experten och därmed mer än 100 % för uppgiften (Ruiz-Primo & Shavelson, 1996). En studie av Acton et al. 1994 visade dock det problematiska i att använda sig av en expertkarta uppgjord av bara en expert. Eftersom begreppskartor mäter kunskapsstrukturer kan också experter ha mycket olika syn på hur kunskapen inom ett givet område är uppbyggd (Ruiz-Primo & Shavelson, 1996). Ett exempel på detta är Richard Feynman som var välkänd för att ha en syn på fysiken som till stora delar skiljde sig från hans samtida men som likväl var rätt (Feynman & Gleick 1994). I bästa fall borde man använda sig av ett slags medeltal av många experter men i praktiken kan detta vara opraktiskt (Ruiz-Primo & Shavelson, 1996).

2.4.4 Tre exempel på bedömningskriterier av begreppskartor

Den första metoden att mäta begreppskartor på presenterades i Novak och Gowins ursprungliga verk, *Learning how to learn* (Novak & Gowin, 1984). Som tidigare nämnts bygger deras begreppskartor starkt på Ausubels kunskapsteori vilken kräver ett hierarkiskt tänkande. Detta återspeglas även i deras mätmetoder. För ett exempel på en karta uppbyggd och poängsatt enligt deras principer, se Bild 3.

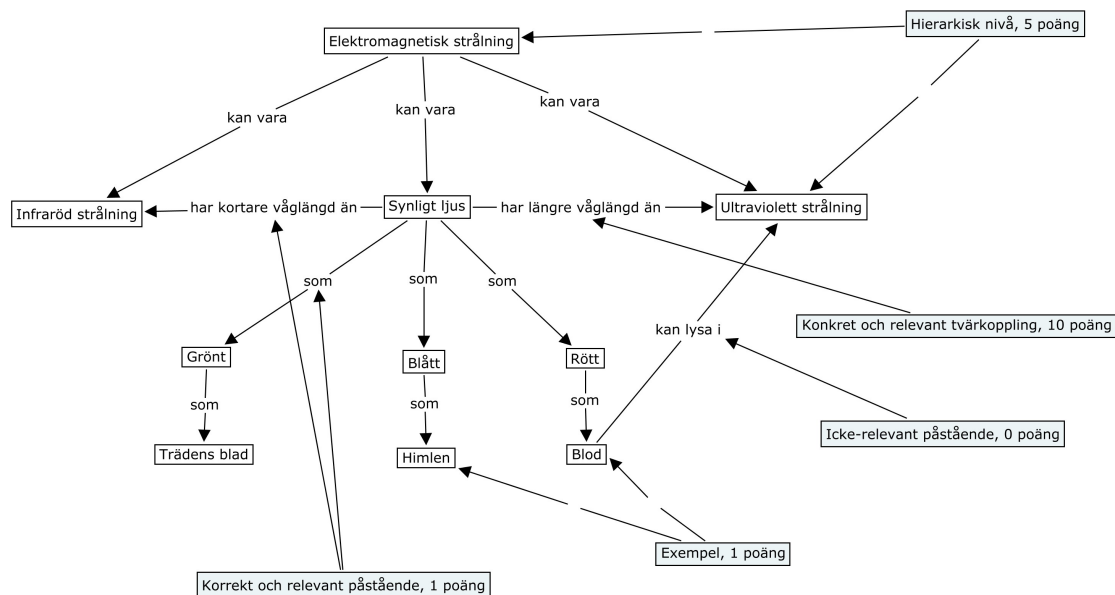


Bild 3 Hierarkisk begreppskarta poängsatt enligt principerna framlagda av Novak och Gowin (Litonius 2012).

Kinchin och Novak ger olika delar av en begreppskarta olika poäng så att man för varje relevant och sant påstående får 1 poäng, för varje hierarkisk nivå 5 poäng, för tvärkopplingar mellan olika begrepp inom kartan 10 poäng ifall den är relevant och korrekt, medan endast 2 poäng om den endast är korrekt, för exempel får man 1 poäng var. (Litonius 2012, Novak & Gowin 1984, sid. 36). Förutom att eleverna konstruerar sina kartor gör man också upp en så kallad expertkarta som används som facit. "Experten" i fråga är i praktiken oftast läraren själv som med sin karta utgör guldstandarden för kunskap i ämnet. Denna expertkarta bedöms enligt samma kriterier som elevernas vilket skapar ett jämförelsetal. Elevernas poängtal jämförs med expertkartans poängtal vilket ger ett procenttal som kan användas både för att bedöma eleven absolut och i förhållande till resten av klassen.

Novak och Gowins strukturella bedömningsform är än idag en av de vanligaste formerna för bedömning av begreppskartor som används för forskning (Ruiz-Primo & Shavelson, 1996) med endast små modifikationer. Som bedömningsgrund är systemet väl-lämpat för att bedöma hierarkiska kartor. Den lägger mycket stor tyngd på tvärkopplingar mellan olika begrepp, kanske onödigt mycket, medan den inte skiljer särskilt mycket mellan begrepp som är korrekta och relevanta och begrepp som endast är korrekta men irrelevanta eller rent ut sagt

det finns ingen skillnad mellan ett begrepp som inte kanske är helt relevant (i lärarens ögon) och ett felaktigt påstående. Det man då går miste om är bland annat sådana begrepp som kanske inte är fullt så viktiga för experten men som för en novis kan vara en viktig brygga till djupare förståelse (Kinchin, Hay, & Adams, 2000).

Den största svagheten i Novak och Gowins bedömningskriterier ligger dock i dess strukturella form. Att tvinga elever in i ett hierarkiskt tänkande kan leda till "gummikartor" (Novak & Gowin, 1984), kartor där den översta delen är helt arbiträrt vald och där det inte är någon större skillnad hur den byggs upp. Vissa ämnesområden är inte lämpade för hierarkiskt tänkande och vissa dynamiska system, som vattnets kretslopp eller fotosyntesen kräver radikalt olika begreppskartor än de som de strukturella systemen kan tillåta (Derbentseva, Safayeni, & Cañas, 2007).

De snäva gränserna Novak och Gowin ställde kan delvis motiveras för vissa ämnesområden där hierarkiska nivåer kommer naturligt, men är onödigt begränsande och arbiträra för många andra ämnen. Gayle Nicoll et al. försökte 2001 skapa ett system att analysera begreppskartor som inte bara skulle tillåta en friare syn på kartans struktur utan som också skulle ge bedömaren mer djuplodande information om elevens kunskapsstrukturer (Nicoll, 2001). De menade att existerande begreppskartor inte gav läraren insikter i hur informationen kom till, hur eleven under kursens lopp så småningom byggde upp sin information. Därför föreslog de att bedömaren skulle använda sig av tre nivåer då begreppskartan analyserades, inte bara huruvida påståendena är korrekta eller inkorrekta, men också hur säker eleven är på sitt påstående samt hur djup information påståendet ger uttryck för (Nicoll, 2001).

Problemen med Nicolls et al. system ifall man försöker tillämpa det på klassrummet hopar sig från första början. Eleverna byggde inte själva upp begreppskartorna, istället intervjuades de i en timme var och begreppskartan byggdes upp utgående från dessa intervjuer (Nicoll, 2001). Redan det här torde göra metoden oanvändbar i ett klassrum med ett par dussin elever. Men också analysskedet är rätt arbetsdrygt. Emedan de flesta begreppskartors bedömningsmetoder siktar på att gå att utföra på ett par minuter, max dryga fem minuter (McClure et al., 1999), verkar den metod Nicoll et al. presenterar ta långt mer än detta. Det måste dock påpekas att Nicoll et al. inte anger hur länge hela processen tog, så man kan bara spekulera.

Ett annat stort problem med den av Nicoll et al. presenterade metoden är godtyckligheten. Korrelationen mellan olika bedömare kom till 0,74 i deras utredning (Nicoll, 2001). Forskarna ansåg detta vara tillräckligt, men andra metoder

har gett korrelationsfaktorer på runt 0,80 (Ruiz-Primo & Shavelson, 1996) (Yin et al., 2005). Att metoden Nicoll et al. rekommenderar har fler faktorer som skapar denna sämre korrelation kan förklara fenomenet men är ingalunda en god sak. Och vare sig bedömarna kommer till samma resultat eller inte visar inte Nicolls et al. studie på huruvida begreppskartan verkligen mäter vad den bör. Till exempel då osäkerhetsfaktorn skulle mätas var det speciellt vissa nyckelfraser som söktes (Nicoll, 2001). Det gör det förvisso lätt för bedömarna att komma till liknande resultat men med tanke på hur olika sätt folk kan ha att uttrycka sig på behöver det inte nödvändigtvis korrelera med någon verklig säkerhetsnivå.

Begreppskartor har visat sig vara användbara för att få en bild av kunskapsstrukturer och för att reda ut missförstånd hos elever. Att försöka använda dem som ett allt för djuplodande mätinstrument tar dock ifrån dem en av deras fördelar, nämligen tidseffektiviteten. Under en kurs har en lärare ett begränsat antal timmar på sig att arbeta med en given elev. Alla verktyg som används måste begründas och deras fördelar måste vägas mot de tidsresurser de kräver (McClure et al., 1999). Nicolls et al. metod kan kanske ha användning för storskalig analys av utbildning och i vilken takt kunskap anammas under en kurs men är inte lämpad för klassrumsanvändning och är därmed inte lämpad för den här studien. Av praktiska skäl kräver klassrumsanvändning att bedömningskriterierna är snabba att tillämpa och ger eleven tillräckligt mycket frihet för att kunna uttrycka sig.

Utmaningen med att hitta ett system för att använda begreppskartor för storskalig eller klassvis bedömning av elevers kunskap är att hitta en balans mellan många varandra uteslutande önskemål. Av praktiska skäl måste både förberedelsen av uppgiften, elevens utförande av den och den senare analysen gå inom rimlig tid. Det här utesluter individuell handledning för varje enskild elev samt allt för komplicerade metoder att analysera resultaten (framtida utveckling av datorprogram kan göra mer avancerade analytiska modeller mer användbara för enskilda lärare). Det skapar också en efterfrågan på effektiva metoder att lära ut begreppskartor åt elever. Förutom det måste det finnas belägg för att bedömningskriterierna är objektiva och att de verkligen mäter något väsentligt. Att lösa dessa problem var det Ruiz-Primo och Shavelson syftade på genom en storskalig undersökning av ett tjugotal undersökningar som använt sig av begreppskartor (Ruiz-Primo & Shavelson, 1996). Senare presenterade de sina egna modeller för hur en begreppskartsundersökning kunde utföras som skulle uppfylla både de pedagogiska och de logistiska kraven.

Den uppgift Ruiz-Primo och Shavelson fann att bäst lämpade sig för praktisk användning var en där eleverna gavs 10 begrepp som hade att göra med det område

de höll på att läsa (Ruiz-Primo & Shavelson, 1997). Dessa begrepp kunde med fördel göras av den som utför forskningen i samråd med läraren eller, ifall det används för rent klassrumsbruk, av läraren ensam eller i samråd med sina kolleger. Dessa 10 begrepp skulle eleverna med penna och papper ordna in i en valfri begreppskarta. De fick använda hur långa länkar som helst i påståendena och det fanns inga krav på att kartan skulle följa någon speciell struktur, som hierarkisk eller cyklisk (Ruiz-Primo & Shavelson, 1997).

För bedömningen gjorde läraren, en expert samt forskarna själva en lista med påståenden baserat på de utvalda begreppen. Dessa samlades och ordnades enligt hur korrekta och relevanta de var, från "inkorrekt" till "utmärkt korrekt" med "irrelevant", "svag korrekt" och "god korrekt" emellan (Ruiz-Primo & Shavelson, 1997). Med hjälp av dessa kunde elevernas påståenden ges poängtal från 0 till 4, med fulla 4 poäng för ett uttömmande, korrekt påstående (Ruiz-Primo & Shavelson, 1997).

För den slutgiltiga bedömningen användes tre metoder. För det första tittade man på elevens absoluta summa av poäng för sina påståenden. För det andra räknade man ut konvergensen, alltså hur många påståenden som matchade de givna påståendena i expertkartan, med andra ord, hur stor del av den totala önskade maximumsumman av poäng eleven fick. Slutligen räknades antalet korrekta påståenden av alla elevens givna påståenden. Den sistnämnda av dessa gav dock mycket mindre pålitliga resultat och kommer därför inte att användas för den här studien (Ruiz-Primo & Shavelson, 1997).

Studien fann att bedömarna var mycket nära varandra i sina resultat (Ruiz-Primo & Shavelson, 1997), vilket också övriga liknande studier visat (Yin et al., 2005) (Ruiz-Primo, Schultz, Li, & Shavelson, 2001), så det är motiverat och befogat att för den här studiens syfte att endast använda en bedömare, trots invändningarna av McClure som nämndes tidigare (McClure et al., 1999).

I den ursprungliga studien gavs eleverna ett flervalsprov innan de skulle göra sina begreppskartor och resultaten jämfördes med varandra (Ruiz-Primo & Shavelson, 1997). Eftersom flervalsprov inte är så vanliga i Finland kommer den här studien istället att använda de reguljära provresultaten eleverna presterar. Detta tjänar vida bättre det tänkta syftet med studien, nämligen att jämföra begreppskartornas resultat med de "vanliga" resultaten, vilket presenteras i nästa kapitel.

3 Studiens syfte, begreppskartan i bruk som bedömningsmetod

Användning av begreppskartor inom utbildningen är en stor del av ämnesdidaktiken på Helsingfors Universitet samt har ett gediget stöd i forskningslitteraturen såväl som inlärningsmetod (Nesbit & Adesope, 2006) som utvärderingsmetod (Ruiz-Primo, 2000). De används dock sällan i skolvärlden. Som studiemetod kräver uppläggandet av en begreppskarta mer aktivt tänkande än enkel läsning (vilket antagligen är orsaken till att den fungerar bättre) och skrivande av franska streck och sammandrag. Skolböckerna är ofta inte uppbyggda för att stöda en koncept- eller begreppsbaserad inläring och begreppen lämnas ofta luddiga och kopplas inte samman. Ifall begreppskartor kunde användas som en utvärderingsmetod skulle eleverna lättare kunna motiveras att arbeta med den metoden, kanske senare också på egen hand, vilket i det långa loppet skulle öka både deras förståelse för ämnen såväl som hjälpa dem minnas stoffet. Båda dessa positiva följder har bland annat Nesbit märkt att är en följd av att studera med hjälp av begreppskartor (Nesbit & Adesope, 2006).

Steget att börja använda sig av begreppskartor som lärare kan många uppleva som långt. Ifall klassen är obekant med begreppskartor måste de först göras bekanta med systemet, diverse studier talar om undervisning som tar mellan 50 och 90 minuter, beroende på stadiet och åldern. Det här betyder åtminstone en om inte två vanliga lektioner och det är sällan en lärare upplever att kurserna brukar ha ett överflöd av sådana. Den totala tidskostnaden kan givetvis bli mindre ifall undervisningen i begreppskartor används som ett undervisningstillfälle i övrigt. I praktiken måste man förbereda sig på att använda fler timmar eftersom frånvaro för några elever alltid kommer att kräva att man går igenom materialet på nytt eller i enrum med eleverna i fråga.

En annan fråga man kan ställa sig ifall man som lärare vill använda begreppskartan som utvärderingsmetod är "Varför?". Många upplever bedömning av begreppskartor som onödigt tidskrävande vilket säkerligen avskräcker från att försöka införa dem som utvärderingsmetod. Syftet med den här studien är att undersöka ifall det går att utveckla en bedömningsmetod för begreppskartor som gör att begreppskartor kan bedömas lika snabbt eller snabbare än till exempel essäbaserade svar. Denna metod jämförs sedan med resultaten av ett sedvanligt prov för att se om resultaten korrelerar. Ifall en stark korrelation finns och arbetsbördan är måttlig är det lättare att ta metoden i bruk.

Begreppskartor som utvärderingsmetod har många fördelar. De tvingar eleverna till ett mer holistiskt sätt att studera och tänka på materialet, inte bara till

annars har svårt att uttrycka sig i text. Den formaliserade stilen på begreppskartorna gör att både djupa insikter samt brister uppdagas lätt, ett exempel på vilket är användningen av begreppet *medium* i studien med grundskolegruppen som mycket få elever kunde koppla ihop med fler än ett par andra begrepp och några lämnade nästan helt okopplade. Det gör det lättare för läraren att senare lappa brister i förståelsen och därför kan begreppskartor anses lämpa sig bättre som mellanförhör, inte som slutprov. Att göra ett förhör för begreppskarta går betydligt snabbare än ett sedvanligt förhör emedan det endast krävs av läraren att begreppen väljs ut på förhand, något som kommer rätt naturligt av ämnesområdet i fråga. Ifall bedömningstiden kan dras ner till ett minimum men korrelationen med sedvanliga provresultat ännu uppehålls kunde begreppskartan användas som ett starkt och snabbt diagnostiskt och utvärderande verktyg.

Naturvetenskap i allmänhet och fysik i synnerhet läses på ett mycket annorlunda sätt i grundskolan och i gymnasiet. Gymnasieundervisningen siktar i praktiken till att förbereda eleven inför studentexamen som till stor del är räkneuppgiftsbaserad, så kallad processuell kunskap (Ruiz-Primo, 2000). Matematikens och problemlösningens andel av gymnasiekursen spelar därmed en betydligt större roll än i grundskolan. I grundskolan är matematiken eleverna har till sitt förfogande ofta bristfällig i fråga om att lösa ekvationer, arbeta med grafer och formler etcetera, så av nöd blir den matematiska delen av fysiken mindre viktig och därmed också den formaliserade problemlösningen. Eftersom begreppskartor inte specifikt mäter problemlösningens förmåga kan man därför anta att begreppskartorna kommer att korrelera bättre med grundskoleelevernas provresultat än med gymnasisternas. Det här är inte i sig ett problem men kan vara värt att ta i beaktande, därför kommer den här studien att undersöka hur dessa två korrelationer förhåller sig till varandra. För grundskolan valdes årskurs 7 som testgrupp eftersom de inte ännu kommit så långt i sina naturvetenskapliga studier och därför kan antas ha lättare att anamma en ny studiemetod. Gymnasiegruppen valdes för att den skolan hade utnyttjat begreppskartor som en inlärningsmetod en längre tid.

4 Material

4.1 Experimentets uppbyggnad

Som testgrupper valdes en klass i gymnasiet och en klass i grundskolan. Grundskolegruppen bestod av 16 elever i årskurs 7, gymnasiegruppen bestod av 13 elever i kurs 6, Ellära, varav tyvärr några resultat måste kasseras vilket beskrivs i

större detalj nedan. Bägge grupper hade fått träning i användning av begreppskartor. Grundskolegruppen hade fått göra begreppskartor tidigare, sammanlagt hade cirka 2 lektioner använts för förberedelse och sedan en lektion med repetition av metoden inför förhøret. Gymnasiegruppen hade dessutom använt begreppskartor som inlärningsmetod och repetitionsmetod under tidigare kurser och kunde därmed antas vara kunniga i metoden. Bägge grupper kan alltså anses ha fått tillräcklig instruktion i att använda begreppskartor, baserat på resultaten i till exempel Maria Ruiz-Primos forskning (Ruiz-Primo et al., 2001).

Uppgiften är baserad på Maria Ruiz-Primos undersökningsmetod som den presenterades i slutet av 1990-talet, med vissa modifikationer. Studien fann att frihet i formen och strukturen på begreppskartor men stark begränsning angående vilka begrepp som får användas gjorde att kartorna kunde bedömas utan större svårigheter men ändå gav insikter i elevens kunskapsstrukturer. (Ruiz-Primo & Shavelson, 1996) (Ruiz-Primo & Shavelson, 1997) (Ruiz-Primo et al., 2001) På basis av detta bestod gruppernas uppgift i att placera in 10 av läraren utvalda begrepp i ett självstrukturerat system. Bedömningen skedde relationellt alltså så att varje påstående gavs ett poängvärde på en skala från 0 till 4 poäng. Så långt följde experimentet den tidigare forskningen. Vissa modifikationer måste göras för att få metoden mera klassrumsvänlig.

4.2 Begreppsval och expertkarta

De ursprungliga studierna var storskaliga, med minst 100 elever och ett antal bedömare i form av lärare, forskare och forskningsassistenter. (Ruiz-Primo et al., 2001) Det här kommer inte att vara fallet i en finsk skola där läraren oftast har mellan 15 och 30 elever att arbeta med och inga assistenter förutom kolleger. I den ursprungliga forskningen skapades både listan på relevanta begrepp och expertkartan, den karta elevernas alster skulle jämföras med, genom konsensus mellan ett antal bedömare. Det här gjordes eftersom jämförelse med individuella expertkartor har funnits vara mycket riskfyllt och ofta kan leda till mycket svag korrelation mellan begreppskartorna och de slutgiltiga vitsorden (Acton, Johnson, & Goldsmith, 1994).

Det finns dock några abers i denna slutsats. I många studier, bland andra Acton et al., bedömdes begreppskartorna delvis strukturellt, alltså enligt hur de är uppbyggda och hurdana länkar man skapar mellan olika begrepp. Ifall poäng endast fås om ens länkar stämmer överens med experternas är det klart att ett medeltal av experter behövs, eller ett antal olika tillåtna strukturer, ifall man inte vill tvinga eleverna till att tänka just som experten. Metoden som här användes, alltså en

bedömning av varje enskilt påstående är kanske begränsat till sin utsträckning men ger eleven en mycket större rörelsefrihet i frågan om att bygga upp sina egna kunskapsstrukturer och ändå få dem godkända. Det måste dock tas i beaktande, att detta delvis går emot den ursprungliga tanken, att då en individs kunskapsnivå stiger kommer dennes tankevärld att börja närma sig experters (Ruiz-Primo & Shavelson, 1996). Det kan tänkas att denna relationella metod förlorar något i prediktiv kraft på grund av detta, men å andra sidan, genom att tillåta eleven att tänka och förstå själv, ger den relationella bedömningen rum för större kreativitet och kräver logiskt tänkande, inte konformism.

Eftersom syftet med den här studien är att se huruvida begreppskartor kan användas som diagnostiskt redskap i en vanlig klass var det nödvändigt att skala bort antalet experter som bestämmer expertkartan samt antalet bedömare. Det senare är problemfritt emedan ett antal studier har visat att begreppskartor kan bedömas konsekvent så till den grad att man kan ge uppgiften åt en enda bedömare och ändå vänta sig ett objektiva resultat, speciellt ifall uttrycken bedöms skilt för sig snarare än holistiskt (Ruiz-Primo et al., 2001). Kvarstår dilemmat med begreppsvalet och expertkartan. För den här studien valdes att läraren/forskaren själv står för dessa. I vardagen är det inte alltid praktiskt att för varje förhör samarbeta med sina kolleger, många av vilka knappast är bekanta med begreppskartor som bedömningsmetod, och tvinga dem att göra en karta man sedan kan jämföra med sitt eget alster. Än mindre är det praktiskt i en skolmiljö att kontakta yttre experter, hur bra det än skulle vara för materialet. Vissa färdigt gjorda begreppskartor kan finnas som är resultatet av samarbete flera experter emellan, men eftersom dessa skulle kräva just det rätta antalet och just de specifika begreppen är det inte heller praktiskt för en lärare att vara helt beroende av sådana. För maximal flexibilitet och tillgänglighet är det bäst om läraren själv kan göra upp listan på använda begrepp samt expertkartan utgående från den egna undervisningen.

I frågan om expertkartan måste vissa försiktighetsåtgärder tas för att den ska vara pålitlig. För det första har erfarenhet från tidigare forskning visat att man har en tendens att rätta sin egen karta med bättre poäng än den förtjänar, till och med om kartan är uppgjord av flera experter (Ruiz-Primo et al., 2001). För det andra är det orimligt att vänta sig lika många länkar och lika djup förståelse för ämnet av grundskoleelever som av deras lärare. För att ta bägge dessa aspekter i beaktande gjordes en expertkarta med de 10 utvalda begreppen (läraren har också här en fördel ifall hen själv valt de 10 begreppen emedan det åtminstone undermedvetet är sannolikt att hen valt sådana som hen upplever att har kopplingar sinsemellan). I

grundskolans fall resulterade detta i en karta med 27 länkar som ansågs vara relevanta och uttömmande. Sådana som skulle vara mindre relevanta lades inte till expertkartan, endast "4-poängs påståenden", alltså totalt 108 poäng. För att motverka lärarens/bedömarens fördelar samt för att sätta det önskade maximala resultatet på en lämplig svårighetsnivå halverades sedan detta resultat, till 54. 54 poäng var därmed det poängtal för samtliga påståenden eleverna skulle få för att få en full 10:a för uppgiften. Ifall de gick över det poängantalet fick de 10+ vilket är vanligt i skolan i fråga. Vitsordet från begreppskartsuppgiften jämfördes sedan med vitsordet från det traditionella provet.

Gymnasiegruppen behandlades på motsvarande sätt med skillnaden att det var deras egna lärare som valde ut begreppen som skulle placeras in i en karta, sammanlagt 11 stycken, medan "expertkartan" skapades av en utomstående. I gymnasisternas fall hade expertkartan sammanlagt 17 länkar, vilket gav en förväntad maximal poängmängd på 34 poäng. I motsats till grundskolegruppen gick ingen av gymnasieeleverna över denna gräns. I samband med att de ritade upp begreppskartan gjorde de ett traditionellt förhör på samma område. På basis av hur de två testen gick fick de ett separat vitsord för vardera prestationen.

Att ge läraren stor frihet och mer bestämmanderätt i bedömningen är vanligt i finska skolor, mer så än i många andra länder. För att kunna använda begreppskartor som en praktisk diagnostisk metod är det viktigt att en enskild lärare kan välja att göra det utan att behöva ty sig till utomstående expertis. Den här studien visar hur lyckat och möjligt det är.

5 Analys

Materialet kommer att analyseras kvalitativt med vissa talande exempel samt kvantitativt för att undersöka hur stark korrelationen är mellan begreppskartans resultat och de traditionella provens. Slutligen kommer resultaten att analyseras för att se huruvida de enskilda resultaten som begreppskartan producerar kan anses stämma överens med de ett traditionellt prov skapar. Materialets ringa storlek gör vittgående statistiska analyser omöjliga men de metoder som används granskar snarast hur enskilda data är kopplade med varandra så den analys som görs lider inte. Däremot är det opraktiskt att ytterligare börja dela upp de två grupperna för att spana in mindre trender som till exempel könsfördelning, huruvida elever med läs- och skrivsvårigheter gynnas av metoden och så vidare. Dyliga undersökningar skulle behöva ett betydligt större elevmaterial att plocka från för att kunna producera ett trovärdigt resultat. Trots dessa begränsningar skall jag försöka dra några försiktiga slutsatser som kan ge idéer till framtida forskning.

5.1 Kartuppgiften som helhet

Överlag klarade de båda grupperna mycket bra av att skapa begreppskartor trots att det i gymnasiegruppen förekom fler formfel och bristfälliga kartor än i grundskolegruppen, en översikt över resultaten kan ses i Bilaga 3. De allmännaste felen var saknade pilar och etiketter, främst hos gymnasisterna. Då länken hade etikett men saknade riktningspil tolkades den ändå som giltig men ifall etikett saknades fanns det inget annat val än att kassera länken helt och hållet. I många fall introducerade också gymnasisterna egna begrepp som i ibland kunde vara användbara och bra att ha med men emedan begränsningen av begreppen är en del av testet måste även de resulterande länkarna förkastas. I några enskilda fall kunde ett tillägg räknas med ifall det extra begreppet tydligt fungerade som länk mellan två eller fler av de givna begreppen. De bristande kartorna ledde till att tre av gymnasisternas kartor helt måste kasseras emedan de saknade alla etiketter vilket tyvärr gjorde den redan betydligt mindre gymnasiegruppen ännu mer fåtalig. Vad dessa bristande begreppskartor på gymnasiesidan berodde på kan man endast spekulera om i efterskott men med tanke på hur väl grundskoleeleverna lyckades hålla sig till de givna parametrarna torde det i framtida experiment gå att åtgärda med hjälp av riktad inlärning av metoden att göra upp begreppskartor.

5.2 Kvalitativ analys

I denna närmast kvalitativa analys kommer de olika gruppernas begreppskartor att synas för att hitta trender och belysande exempel på elevernas tankemodeller. Förutom att försöka hitta underliggande trender kommer en analys av elevernas tankevärldar också att göras genom att antalet utgående och inkommande länkar till diverse begrepp beräknas. Här tas inte de enskilda länkarnas riktning eller faktuell relevans i beaktande eftersom det här är för att mäta hur viktigt det här begreppet är för eleven i fråga, när hon bygger upp sin kunskapsvärld. Den här problematiken presenterades av Kinchin et al. som menade att en analys som endast tar i beaktande de korrekta länkarna mycket väl kan gå miste om strukturer som är relevanta för eleven i fråga. (Kinchin et al., 2000) Också felaktiga kopplingar kan hjälpa att temporärt koppla samman begrepp där den korrekta kopplingen än så länge är i ett formativt stadium.

Kinchin & Hay identifierade tre olika begreppskartsstrukturer, fritt översatta till kedja, nav och nätverk (chain, spoke och network). (Kinchin et al., 2000) Naven är mycket centraliserade strukturer där allt utgår från ett nyckelbegrepp men som inte är direkt kopplade till varandra. En kedja följer en logisk struktur där ett begrepp följer på ett annat. Kedjan ger ett falskt intryck av hierarki begreppen emellan. Nätverket är just vad det låter som, ett väl sammanvävt nätverk av

kopplingar mellan de olika begreppen. Navstrukturen tillåter ny information att införas utan att det stör helheten överhuvudtaget med enda aber att läraren måste komma ihåg att koppla det nya begreppet till nyckelbegreppet. En kedjestruktur kan vara svårare att omvandla. Ifall det finns ett tydligt avbrott i kedjan eller ifall nya begrepp läggs till i slutet är det inte svårt, men det kan vara svårt för en elev som skapat en kedjestruktur att integrera ny information i mitten av kedjan. Ny information som läggs till i början kan vara så omvälvande att eleven snarare förkastar informationen. Nätverket är flexibelt när det gäller att koppla in ny information men den måste rotas väl i många begrepp för att passa in (Kinchin et al., 2000). Några exempel på de olika strukturerna presenteras nedan.

I verkligheten är det sällan någon begreppskarta är renodlat någon av dessa tre strukturer (förutom nätverk) så vissa tolkningar måste alltid göras. En snabb länkanalys hjälper en ofta att luska ut vad det snarast är frågan om. I den här studien har kartor där den absolut största delen av begreppen har en eller två länkar till eller från sig och där inget begrepp har mer än tre länkar räknats till kedjestrukturer. Navstrukturer har tolkats som de som har en eller ett par begrepp med betydligt fler länkar än de andra och de andra har en eller två. Ifall fördelningen är jämnare och många begrepp har fyra eller fler länkar börjar kartan vara ett nätverk. Dessa är kvalitativa regler, inte stenhårda klarhuggna kategorier och speciellt navkartorna har en tendens att närma sig nätverken om de blir mer komplicerade. I många fall är det mycket svårt att säga var skillnaden mellan ett mycket centraliserad nätverk och ett hoptrasslat nav går.

Genom att analysera länkarnas antal kan man ytterligare få fram vilka saker som för eleverna är nyckelbegreppen. För en lärare är det viktig information som kan berätta vilka begrepp man skall sätta speciell vikt på eftersom de är kritiska för elevernas helhetsbild. Ifall eleverna gång på gång har samma begrepp som sina kärnnoder är det för det första viktigt att se till att deras bild av dem är korrekt och för det andra kan det vara tacksamt att se till att nytt material kopplas just till dessa kärnbegrepp. Å ena sidan kan det här leda till en onödigt stark navkarta men å andra sidan kan det vara till hjälp för nybörjare för att de överhuvudtaget ska kunna inordna sin kunskap i något slags system. Om däremot expertkartan har en mycket annorlunda struktur med helt andra relevanta nyckelbegrepp kan det vara värt för läraren att fundera på vad det kan bero på och också satsa extra på att ge mer vikt åt onödigt försummade begrepp. Typiska exempel på sådana är *medium* och *resonans* på grundskolestadiet, vilka båda visade sig vara begrepp många hade mycket svårt att koppla ihop med det övriga materialet.

Det finns fler begreppskartor på grundskolesidan och en större variation mellan dem, så en större del av analysen av enskilda kartor kommer att behandla dem. Grundskoleelevernas kartor ger också en bra möjlighet till att presentera poängsättningsprinciperna i praktiken. I korthet delades alla begrepp upp i fem kategorier enligt hur korrekta och relevanta de är och poängsattes därefter. Det är svårt att med det begränsade materialet hitta något enskilt begreppspar som skulle ha producerat länkar av alla kategorier, men åtminstone kan en viss trappstegseffekt noteras:

0 poäng, felaktigt påstående. "Frekvens \rightarrow påverkar höjden på \rightarrow Amplitud". Påståendet är att hur snabbt en vågrörelse varierar kommer att påverka höjden på vågorna. Även om det stämmer med till exempel pendlar att större hastighet på pendeln leder till större amplitud i svängningen kommer frekvensen på pendelns rörelse inte att påverkas eftersom den ökade amplituden leder till att pendeln måste röra sig längre. Det här är ett grundläggande missförstånd som stämmer väl överens med en vardaglig syn på fenomenet svängningar, men som inte stämmer överens med en närmare analys av fenomenet.

1 poäng, irrelevant men korrekt påstående. Dessa är bland de svåraste att hitta ett rent exempel på emedan de flesta elever antingen har något som är fel eller så påståenden som åtminstone delvis är relevanta. De vanligaste länkarna som klassades i den här kategorin är de som är så vaga att de inte riktigt säger något, som "Volym \rightarrow påverkas av \rightarrow Svängningsrörelse". Det är tekniskt sett korrekt men anger inget om hurdant förhållandet är, på vilket sätt de är sammankopplade och så vidare. De irrelevanta påståendena skiljer sig oftast mycket lite från länkar utan etiketter till informationsvärde. De betecknar att två begrepp på något sätt är sammanlänkade men hur framgår inte.

2 poäng, bristfälligt men korrekt påstående: "Frekvens \rightarrow skapar \rightarrow Resonans". Dessa påståenden visar redan att eleven har ett visst hum om den bakomliggande kopplingen mellan två begrepp men detaljerna saknas. Resonans mellan två oscillatorer uppstår på grund av deras frekvenser men på vilket sätt framgår inte av den här länken så den måste anses bristfällig.

3 poäng, korrekt påstående: "Resonans \rightarrow uppstår ifall två eller flera oscillatorer har samma \rightarrow Frekvens". I jämförelse med det tidigare påståendet har den här eleven visat sig förstå förhållandet mellan frekvens och resonans i mycket högre grad. Påståendet är korrekt såtillvida att två oscillatorer med samma frekvens på svängningarna kommer att resonera med varandra. För elever på högre nivå kunde det här klassas som ett bristfälligt påstående men för nybörjare är detta redan ett hyggligt exempel på kunskap.

4 poäng, uttömmande och korrekt påstående: "Resonans \rightarrow uppstår då 2 st har samma eller jämn multipel av \rightarrow Frekvens". Denna elev har bemästrat kursens innehåll vad angår förhållandet mellan frekvens och resonans. För att resonans skall uppstå krävs inte att frekvensen är fullständigt lika, det räcker med att den ena oscillatorns frekvens är en jämn multipel av den andra. Med tanke på det begränsade mediet begreppskartan är för långa utredningar kan man inte vänta sig en mer utförlig beskrivning av förhållandet mellan dessa två begrepp.

Denna poängsättningsprincip fungerade som riktlinje för både grundskole- och gymnasiegruppens kartor, nedan följer en närmare kvalitativ diskussion om de två gruppernas resultat. Angående strukturerna kan sägas att av grundskoleelevernas kartor kan 6 klassas som kedjor, 5 som nätverk, 4 som navar och en är svår att bedöma. För gymnasiet är saldot 6 kedjor, 2 nätverk, en nav och en blandform. Märk väl att skillnaden på nätverk och navar ibland kan vara hårfin, därav följer blandformen. Kedjor hör till de svagaste begreppskartorna och i gymnasiefallet var över hälften sådana, medan drygt en tredjedel av kartorna från grundskolegruppen var av den typen.

5.2.1 Begreppskartorna från grundskolegruppen

Området handlade om vågrörelser så det är att vänta sig att begreppet "Vågrörelse" kommer att vara ett nyckelbegrepp för många, nära följt av "Svängningsrörelse", eftersom så gott som samtliga övriga begrepp mycket enkelt går att koppla direkt till dessa. Expertkartan hade "Vågrörelse" som det överlägset centralaste begreppet med 10 länkar till eller från, med "Svängningsrörelse" som god tvåa med 7. Det visade sig att elevernas kartor avspeglade det här mycket väl. 9 av de 16 eleverna hade "Vågrörelse" som nyckelbegrepp och endast fyra elever hade varken "Vågrörelse" eller "Svängningsrörelse" som den starkaste noden. Mycket talande är också att de fyra som inte hade någondera hörde till de svagaste i gruppen, med två underkända kartor (en av vilken endast innehöll två av de tio begreppen), en med vitsordet 5+ och en med 6. Förutom eleven som fick 6 i kartuppgiften men fick 7 $\frac{1}{2}$ i provet, presterade de också mycket svagt i det traditionella provet, med två underkända och en med 5 $\frac{1}{2}$. Trots att det fanns en del elever av de som hade de två begreppen som nyckelbegrepp och som ändå gjorde svagt ifrån sig finns det åtminstone ett gott indicium på att ifall man inte ens lyckas koppla dessa två saker till de andra begreppen har man inte riktigt greppat kursen och kommer därmed att få dåligt resultat oberoende av mätmetoden. Eftersom materialet så starkt bygger på dessa kärnbegrepp är de naturliga utgångspunkter för så gott som hela det övriga materialet

Resultaten från grundskolesidan var mycket varierande men vissa trender kan ses. Det fanns elever som antingen inte anammat begreppskartsstrukturen eller som hade en mycket lösryckt och splittrad kunskapsbas. En sådan elev var P vars begreppskarta presenteras i Bild 4.

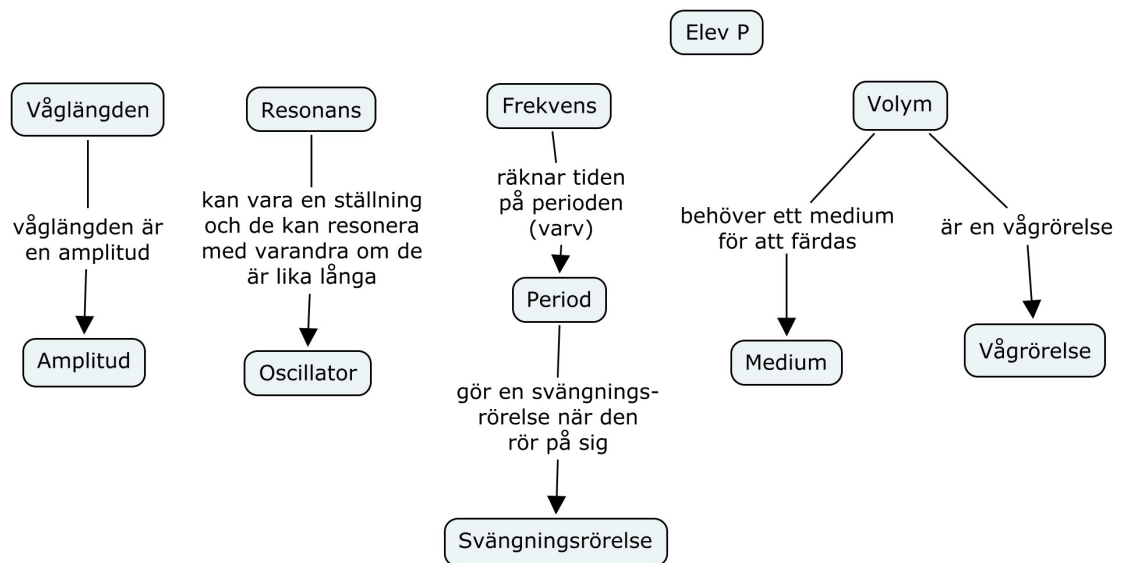


Bild 4 Elev P från grundskolan. Eleven har inte anammat begreppskartans struktur eller så är elevens kunskapsbas mycket lösryckt och splittrad.

P:s karta består av lösryckta fraser och kan egentligen inte ens uppfattas som en begreppskarta. Länkarna skapar bristfälliga eller fullständiga nonsens-satser, som till exempel "Våglängden → våglängden är en amplitud → Amplitud" eller "Volym → behöver ett medium för att färdas → Medium". Mycket konkreta minnesbilder från demonstrationer som "Resonans → kan vara en ställning och de kan resonera med varandra om de är lika långa → Oscillator" som verkar hänvisa till en demonstration som gjordes under lektionen. Demonstrationen har fastnat i huvudet, likaså orden resonans och oscillator, men kartan demonstrerar ingen större förståelse för fenomenet.

Vitsord: karta 4, förhör 5½

Andra elever uppvisade god förståelse av det inlärd. En sådan elev var C, vars begreppskarta presenteras i Bild 5.

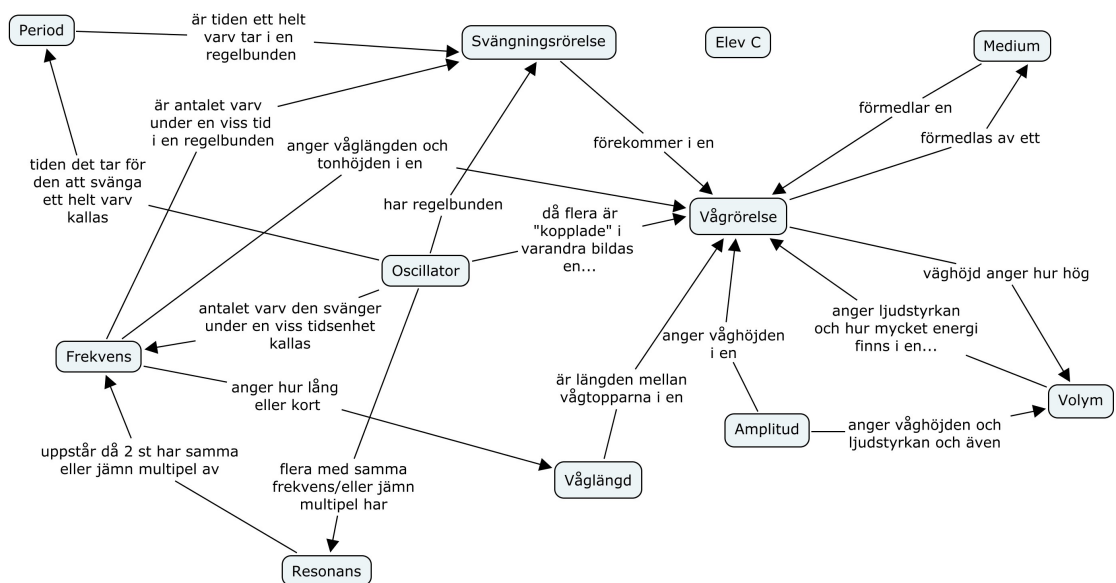


Bild 5 Elev C från grundskolan. Begreppskartan har en stark och relativt välbalanserad nätverksstruktur där majoriteten av länkarna är mycket välformulerade.

C:s karta utgjordes av ett bra nätverk och var den mest utförliga kartan i gruppen. Länkarna visar på en för kursen fullständig förståelse för fenomenen, som kopplingen mellan oscillator och vågrörelse: "Oscillator → då flera är 'kopplade' i varandra bildas en... → Vågrörelse" samt kopplingen mellan de vardagliga och mer teoretiska delarna av begrepp, som "Frekvens → anger våglängden och tonhöjden i en → Vågrörelse", där det i frågan om ljud kommer fram både det vi kan höra (tonhöjden) samt det vi vet om ljudvågornas längder (våglängden). Också här kommer experimentet med resonatorn fram genom "Oscillator → flera med samma frekvens/eller jämn multipel har → Resonans" samt "Resonans → uppstår då 2 st har samma eller jämn multipel av → Frekvens" som visar en betydligt större förståelse för hur resonans uppstår och vad det beror på än i fallet P.

Vitsord: karta 10+, förhör 9+

Många elever befann sig någonstans emellan P:s och C:s nivå. Ofta hade dessa elever skapat kedjestrukturer som i Bild 6.

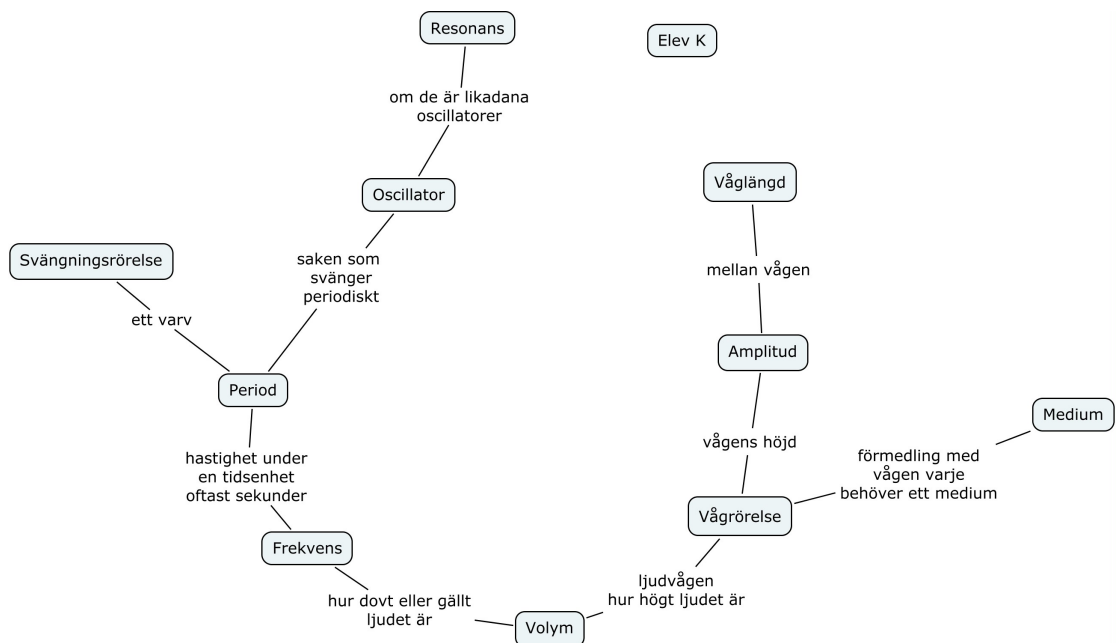


Bild 6 Elev K från grundskolan. Typisk kedjestruktur där de flesta begreppen har en eller två länkar kopplade till sig.

I den kedja K ritat upp saknas för det första riktningspilar för begreppen, men största bristen ligger i strukturen. Nästan inga begrepp, förutom *period* och *vågrörelse* har mer än två länkar till sig. Det är alltså en klassisk kedjestruktur. Kopplingarna är inte alltid i sig så bristfälliga, men eleven har inte utnyttjat begreppskartans fulla potential, det är snarast lösryckta fraser som kopplar två begrepp, som "Oscillator - om de är likadana oscillatorer - Resonans" och "Vågrörelse - förmedling med vågen behöver ett medium - Medium". Vissa tydliga missförstånd kan ses, som att ljudets volym anger hur dovt eller gällt det är samt att amplituden är kopplad till våglängden.

Vitsord: karta 6+, förhör 6+

I vissa fall kunde kartans struktur i sig vara bra även om länkarna var svaga och innehöll många missförstånd. Ett exempel på en sådan karta ses i Bild 7.

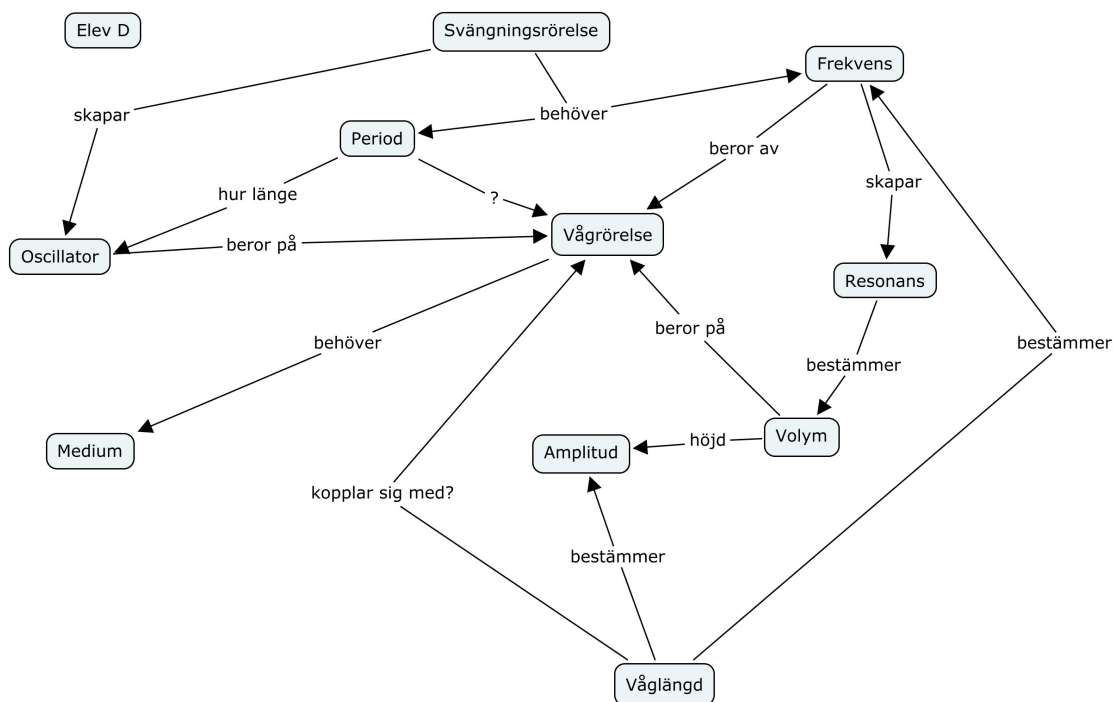


Bild 7 Elev D från grundskolan. Kartans struktur är i sig bra men länkarna är svaga och innehåller många missförstånd.

Eleven D har skapat ett i princip hyfsat, om något centraliserat, nätverk men länkarna är ytterst svaga. Det här är ett bra exempel på hur kartorna kunde se ut för många av de svagare eleverna, nämligen länkar som endast är ett ord eller två och som, trots att de stämmer inte egentligen ger så mycket tilläggsinfo. Länkar som "bestämmer", "behöver", "beror på" och så vidare kopplar samman en stor del av begreppen, som i fallet "Volym → beror på → Vågrörelse" eller "Frekvens → skapar → Resonans". På grund av begreppskartans poängsättningsprincip ger dessa mellan 1 och 3 poäng, men det är svårt att veta hur bra förståelse som egentligen ligger bakom ett kort uttryck som "Vågrörelse → behöver → Medium". I frågan om dessa blir bedömningen nästan arbiträr, eftersom det kan vara ett desperat försök att på något vis koppla ett fullständigt okänt ord till vågrörelser och "behöver" är ganska säkert. Eller så kan det vara ett mycket förkortat uttryck för en djup insikt i att alla vågrörelse behöver ett medium som oscillerar. Länken i sig ger inte tillräckligt med information.

Vitsord: 6½, 7

5.2.2 Begreppskartorna i gymnasiet

Kursen var ellära och det relevanta avsnittet cirkulerade mycket kring elektrisk effekt. Expertkartan byggdes upp med *elström* som starkaste begrepp och *källspänning* som god tvåa. Liksom i fallet med grundskolan hade en majoritet av eleverna någon av dessa som starkaste eller likvärdigt starkaste nod men nu var marginalen mycket mindre, hela fyra av tio hade andra begrepp som de viktigaste och i så gott som alla av dessa var *resistans* det mest länkade begreppet, något som

skall granskas närmare nedan då elev K från gymnasiet diskuteras. Många av gymnasisterna hade lagt in en hel del egna begrepp samt länkar utan etiketter. Till stor del slopades dessa i bedömningen vilket också ledde till att några hel förkastades för den här studien. Sådana länkar som går via en hel del andra begrepp har ändå tagits med i de datoriserade begreppskartorna som streckade linjer. I vissa fall har ett utomstående begrepp helt enkelt inkluderats som en del av en länk mellan två begrepp ifall det går att dra en direkt koppling genom det, som i fallet med elev H, där ett länkpar "Elström \rightarrow kan utföra: \rightarrow Arbete \rightarrow /tid \rightarrow Effekt" tolkats som Elström \rightarrow kan utföra arbete/tid \rightarrow Effekt".

I det här skedet måste tilläggas att en kontroll av starkaste noder i frågan om gymnasiekartorna har betydligt mindre relevans än i grundskolekartorna emedan en ansevärt större andel av gymnasiekartorna kan klassificeras som kedjor och de starkaste noderna har två eller tre länkar tillsammans med många andra noder. Detta är fallet i hälften av gymnasiekartorna. Därmed går det inte att dra lika långtgående slutsatser om elevens kunskapsmängd endast utgående från detta som i grundskolegruppens fall. Medeltalet för de som inte hade någondera av de två begreppen expertkartan byggdes runt, *elström* eller *källspänning*, som sina viktigaste noder är ungefär $6\frac{1}{2}$ medan medeltalet för hela gruppens begreppskartor är 7, så ingen större skillnad där. Här är det en verkligt stor skillnad med grundskolegruppen där medeltalet för hela gruppen likaså är 7 men där medeltalet för de som inte använde de vanligaste begreppen som sina kärnbegrepp är 5-, mer än en skillnad på två vitsord alltså.

Många av eleverna i gymnasiet representerade en mycket matematisk syn på materialet, som kan ses i Bild 8:

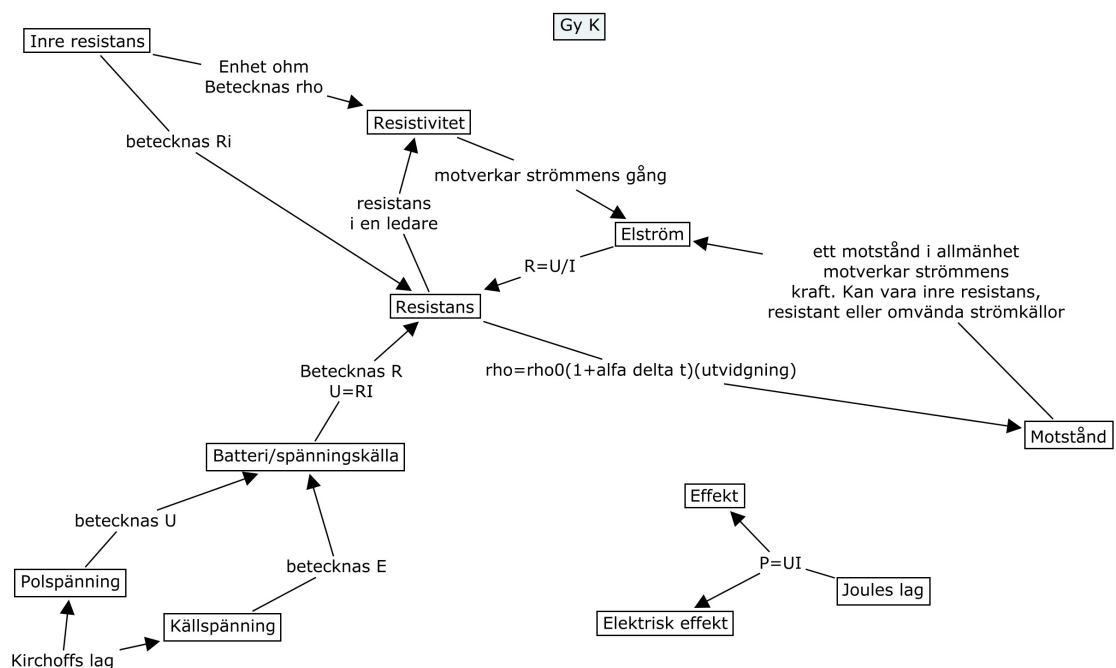


Bild 8 Elev K från gymnasiet. Kartan i sig är mycket starkt centrerad runt "Resistans" som begrepp medan länkarna visar på en förståelse som främst grundar sig på matematisk problemlösning och formler.

Eleven K:s syn på materialet är från grunden baserat på uträkningar och formler. Strukturen är uppbyggd med stark fokus på resistansen och hur den matematiskt är kopplad till diverse andra begrepp och fenomen som motstånd, resistivitet, elström och spänningskällor. Speciellt intressant är att eleven till exempel inte har en direkt koppling mellan "Spänningskälla" och "Elström" utan också det här förhållandet går via "Resistans". En stor del av länkarna är irrelevanta för egentlig förståelse, som "Polspänning \rightarrow betecknas $U \rightarrow$ Batteri/spänningskälla" samt "Källspänning \rightarrow betecknas $E \rightarrow$ Batteri/spänningskälla". Påståendena är i sig korrekta och nödvändiga för att kunna utföra vissa räkneuppgifter, men på ett djupare plan är beteckningar en bisak, inte en huvudsak. Talande för den här matematiska, lösryckta synen är även att begreppen "Effekt", "Elektrisk effekt" och "Joules lag" är sammankopplade med formeln " $P = UI$ " men inte på något vis kopplat till resten av helheten (originalkartan har pilar till "Batteri/källspänning" men dessa har ingen etikett och har därför räknats bort). Eleven verkar ha en förståelse för materialet som helt och hållet är uppbyggd enligt lösning av vissa matematiska problem utan större koppling till vad fenomenen innebär i verkligheten. Med tanke på det är det inte förvånansvärt att resultatet blev bättre i det sedvanliga provet än i begreppskartan där problemlösning och utantillinläring kan ha en starkare inverkan på vitsordet än helhetsbilden.

Vitsord: 7-, 8

Eleven K gav uttryck för en måhända ensidig men trots det gedigen förståelse för materialet. Tyvärr var så inte fallet för många av gymnasieeleverna som inte visade

någon som helst stark förståelse för kursens innehåll. Ett exempel på detta är elev B vars begreppskarta presenteras i Bild 9:

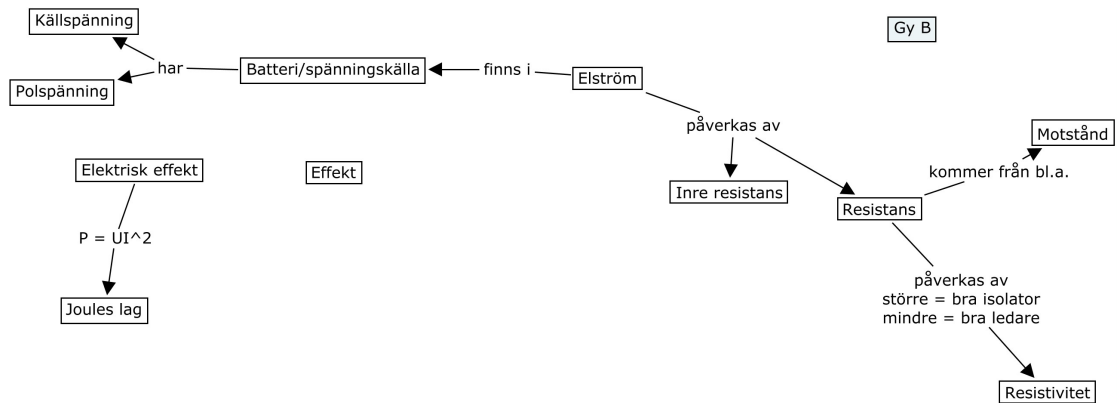


Bild 9 Elev B från gymnasiet. Kartan är svag både till innehåll och struktur. Få noder är kopplade till mer än två andra noder, vissa svävar fullständigt fritt och de felaktiga länkarnas antal är högt.

Begreppskartan elev B ritat upp visar på fundamentala brister i förståelsen inom området och är därmed ett ypperligt exempel på en av begreppskartornas styrkor, nämligen att blottlägga sådant som blivit fel lärt, kanske långt innan den aktuella kursen. Uttryck som "Elström → finns i → Batteri/spänningskälla" och "Resistans → kommer från bl.a. → Motstånd" ger en bild av ström och resistans som konkreta föremål som kan hittas i diverse elektriska komponenter, måhända en följd av en allt för långt driven liknelse på grundskolan då elektriska kretsar först diskuteras. Det här är dock endast problem på en teoretisk nivå, för att lösa problem eller uppgifter är det ingen skillnad ifall elström är något som ligger på lur i ett batteri för att sedan springa iväg längs ledningarna så fort chansen ges, U kommer fortfarande att vara lika med RI. Trots att det traditionella förhållandet som gjordes i samband med begreppskartan inte innehöll några sådana räkneuppgifter var uppgifterna i förhållandet enklare att lära sig utantill (se Bilaga 2).

Vitsord: 5, 7½

5.3 Kvantitativ analys, en jämförelse mellan begreppskartan och traditionella prov som mätmetod

Innan någon fortsatt diskussion om hur begreppskartan klarade sig som mätmetod måste några ord sägas om jämförelsens grunder. I de tre följande analysmetoderna antas det traditionella provet läraren i fråga brukar göra vara det normala och begreppskartan jämförs med detta. En kursrisk analys av de bägge lärarnas vanliga prov görs i hopp om att förklara de insamlade uppgifterna, men någon kritik av dessa görs ej. Orsaken är att de prov lärarna i fråga använder sig av i vardagen med sina klasser på gott och ont är dagens accepterade standard, ett resultat av deras erfarenheter, utbildning och läroplanen. Att försöka gå in på en djupare analys av huruvida det är de traditionella proven eller begreppskartan eller måhända någon

område. Ingen utvärderingsmetod kan med säkerhet sägas vara den allena saliggörande och allomfattande men för den här studien antas den som vanligtvis används vara den man bör jämföra med, en slags guldstandard för utvärdering, den en stor del av elevernas vitsord är baserad på och därmed i längden den metod som ger samhället en bild av vad eleven kan tänkas kunna i ett ämne.

5.3.1 Svårighetsgrad och medeltal

De enskilda elevernas resultat presenteras i detalj i Bilaga 3. Tabellen i Bilaga 3 ger vid första överblick ett rätt hyfsat samband vad gäller svårighetsgraden för de två olika uppgifterna. Förvisso har vissa elever presterat bättre med än den ena metoden och än den andra, men medeltalen för de två metoderna är relativt nära varandra i bägge fallen. För grundskoleeleverna är medeltalet för det traditionella provet $\bar{x}_{prov} = 7,14$ och för begreppskartan $\bar{x}_{karta} = 6,88$ ($\Delta x = 0,26$) medan ditoresultat för gymnasiet är $\bar{x}_{prov} = 7,48$ och $\bar{x}_{karta} = 6,98$ ($\Delta x = 0,50$), i bägge fallen en skillnad som trots allt faller väl innanför acceptabla gränser med tanke på att grupperna använt begreppskartor betydligt mindre än de skrivit traditionella prov och förhör. Ytterligare erfarenhet och övning med kartorna kunde väl tänkas höja begreppskartornas medeltal till ungefär detsamma eller mer än de traditionella provens.

Att medeltalen är så pass nära varandra och kunde tänkas bli närmare med lite mer erfarenhet från elevernas sida kan ses som ett första tecken på att begreppskartan som utvärderingsmetod är giltig och att svårighetsgraden kan antas vara lämplig. Medeltalet i det traditionella provet ger en översikt över kunskapsnivån hos de två grupperna och skulle medeltalet ha varit mycket annorlunda skulle man kunna dra slutsatsen att begreppskartsuppgiften antingen är för lätt, för svår, för obegriplig eller bedömd med felaktig skala. Valet att lägga den väntade poängmängden för expertkartan till två gånger antalet expertlänkar är till stor del arbiträr, liksom valet att i ett prov lägga gränsen till godkänt på 30, 40 eller 50 procent, men visade sig ge en lämplig svårighetsgrad för eleverna.

5.3.2 Korrelation

Ett vanligt sätt att mäta hur väl två dataset stämmer överens med varandra är genom att mäta deras korrelation till varandra. I det här fallet har vi två värden för varje elev, vitsordet eleven skulle ha fått ifall begreppskartan skulle vara den enda mätmetoden och vitsordet eleven fick i det traditionella provet, den så kallade guldstandard. I det här fallet väntar vi oss en lineär, positiv kausalitet. Ifall man skall kunna ersätta det traditionella provet med en begreppskarta förväntar vi oss att en elev som gör bra ifrån sig i det traditionella provet också gör bra ifrån sig i

här stämmer kan man använda sig av Pearsons korrelationsekvation som ger korrelationen r mellan två uppsättningar data, X och Y:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

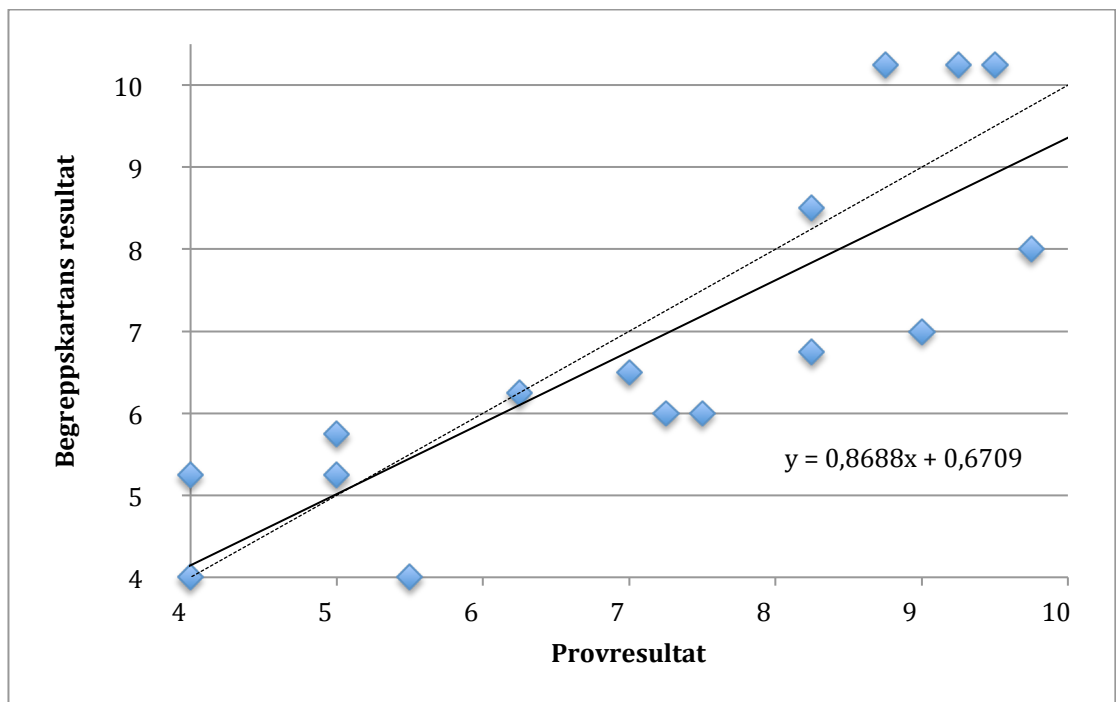
Formeln öppnar sig inte enkelt för en vanlig läsare, men i korthet kan sägas att formeln granskar alla talpar för de två uppsättningarna data, X och Y. Den övre raden i ekvationen (täljaren) jämför varje enskilt värde i gruppen med gruppens medeltal och multiplicerar resultatet. Detta ger ett positivt värde ifall både x- och y-värdet är högre än eller lägre än sin grupps medeltal, men ett negativt värde ifall den ena är högre och den andra är lägre. Den nedre raden (nämnaren) ger alltid ett positivt värde där de absoluta avvikelserna från medeltalet för de båda serierna jämförs med varandra. Matematiken är något krånglig och behöver inte för den här studiens syfte granskas i större detalj, men en viss diskussion om följderna av Pearsons korrelationsformel är av vikt för att förstå den senare analysen.

Resultatet för r är ett tal mellan -1 och 1, där 1 betyder perfekt korrelation, alltså att varje gång x ökar kommer också y att öka, medan -1 betyder det motsatta, alltså att varje gång x ökar kommer y att minska. Ett resultat på 0 betyder att formeln inte funnit någon koppling mellan de två uppsättningarna data. I det här fallet skulle en stark korrelation kunna tolkas som att både det traditionella provet och begreppskartan mäter något liknande. Eftersom traditionella prov och begreppskartor skiljer sig från varandra rätt mycket skulle det här ge bägge metoderna ökad trovärdighet, eftersom två så vitt skilda metoder kommer till liknande resultat. I större skala skulle det ge hopp om att bägge faktiskt mäter något relevant om vad eleven i fråga kan.

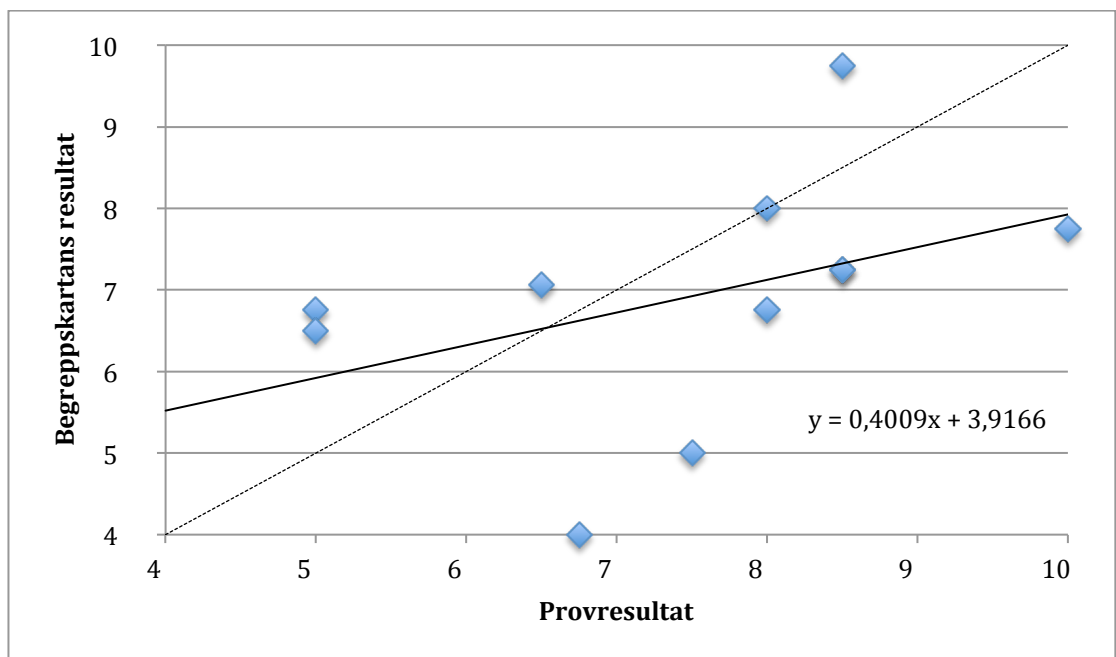
Ifall korrelationen är svag eller till och med negativ är det svårare att dra direkta slutsatser om provens giltighet. Det kan vara att den ena eller båda metoderna är bristfälliga i sin mätmetod eller så kan det vara att de mäter mycket olika saker vilket därför kan förväntas leda till mycket olika resultat. Också om korrelationen är stark bör man dock vara försiktig med att dra allt för långt gående slutsatser.

Då resultaten för de två mätningarna presenteras grafiskt visar sig en klar skillnad mellan grundskolan och gymnasiet. Resultaten kan ses nedan i Graf 1 och 2. Den streckade linjen representerar punkternas riktlinje ifall begreppskartan och provresultatet gav identiska resultat, alltså att eleverna i medeltal gjorde precis lika ifrån sig i bägge testen. Den svarta linjen representerar den verkliga trenden av resultaten. Korrelationen anger hur mycket de enskilda punkterna i medeltal skiljer

sig från den streckade linjen. Ekvationen anger lutningsgraden på den verkliga trenden och visar en indikation på hur väl de två värdena korrelerar med varandra.



Graf 1 Resultat av provet och begreppskartan för grundskolan. Den svarta linjen är trenden med lutningsvinkel 0,869 medan den streckade linjen anger trenden ifall resultaten för testen var identiska.



Graf 2 Resultat för provet och begreppskartan för gymnasiet. Den svarta linjen är trenden med lutningsvinkel 0,401 medan den streckade linjen anger trenden ifall resultaten för testen var identiska.

Korrelationen för de olika grupperna kan räknas ut till 0,825 för grundskolan och 0,412 för gymnasiet, alltså ungefär märkbart mindre.

Korrelationen mellan begreppskartan och provet är betydligt starkare för grundskolan än för gymnasiet. För grundskolan är korrelationen så hög att man kunde tänka sig att de två faktiskt kunde kompensera varandra, medan

gymnasieresultatet endast ger en mycket lös koppling mellan värdena. Det finns en teoretisk förklaring till detta, trots att man måste vara försiktig med att dra allt för långt gående slutsatser utgående från enkel korrelation, vilket presenteras i följande avsnitt.

Fysiken i grundskolan lider genomgående av att eleverna har mycket bristfällig kunskap i matematik. Matematiken som skulle behövas för att närma sig fysiken via formler och därmed göra den mer exakt och problemlösande kommer långt efter kurserna i fysik, speciellt på årskurs 7 som studerades nu. Det här leder till att kursmaterialet är uppbyggt mer baserat på begrepp och deras förhållanden till varandra än på att kunna räkna specifika uppgifter. Därmed är stoffet mycket väl lämpat för begreppskartor. Gymnasiet är å andra sidan mycket mer matematikbaserat emedan kurserna strävar efter att förbereda eleverna för studentexamen som överväldigande bygger på räkneuppgifter. Matematikkunskaperna hos de som läser fysik i gymnasiet antas vara betydligt starkare och kursen är byggd därefter. Det leder till att fokus har en tendens att ligga på formeljonglering snarare än en heltäckande överblick över området. Detta kunde märkas genom att många av gymnasisternas begreppslänkar gick i stil med "räknas ut med", "är formeln för" och så vidare.

Begreppskartor som utvärderingsmetod mäter en persons tankevärld, hur olika begrepp hänger ihop och huruvida en fungerande helhet med många tvärkopplingar kan byggas upp enligt den. En bra förståelse för ett område leder till god fingertoppskänsla för hur ett problem är uppbyggt och måhända till en föräning om hur en lösning kunde se ut och vad resultatet kunde tänkas vara men leder inte direkt till en förmåga att matematiskt lösa problemen. Grundskolegruppens begreppskartor korrelerade starkt med deras provresultat eftersom båda i grund kan antas mäta ungefär samma sak, medan gymnasiegruppens skiljde sig storligen av den motsatta orsaken. Gymnasiet verkar satsa mer på problemlösning och räkning än på en heltäckande förståelse, vilket ger mycket svagare och mycket mer varierande resultat då deras vanliga prov jämförs med begreppskartor.

5.3.3 Överensstämmelse med de traditionella förhören

Korrelationsanalysen verkar ge vid handen att begreppskartan fungerar som ett potentiellt mycket bra kompenserande utvärderingsverktyg åtminstone i grundskolan. En närmare analys visar dock att den synen är aningen överoptimistisk. Korrelation mäter endast styrkan i kopplingen mellan två variabler, inte hur väl de verkligen stämmer överens med varandra. Som exempel kan nämnas ett värde som kan väntas vara lineärt. I så fall kommer korrelationen att vara mycket hög ifall en annan mätserie också ger ett lineärt värde med något liknande

lutningsvinkel, trots att de enskilda värdena kan variera storligen. (Bland & Altman, 1986)

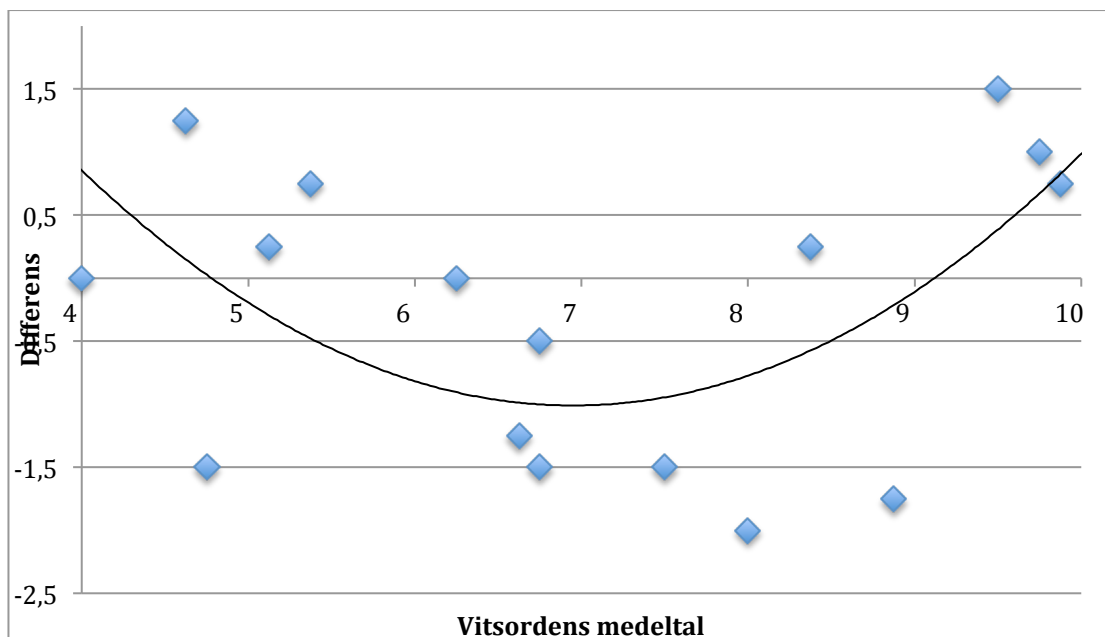
Den klassiska frasen om att korrelation inte indikerar kausalitet kan tillämpas också i det här fallet. Det att två uppsättningar data ger en liknande trendlinje betyder att resultaten korrelerar starkt men inte att de i fall av enskilda elevers testresultat kan anses likvärdiga. Eftersom korrelation räknas genom att bland annat dividera resultaten med standardavvikelsen kommer en större spridning bland resultaten alltid att leda till en större korrelation än ifall man mätt ett snävare område. Det här gör att man omedvetet kan få bättre resultat än vad som egentligen är fallet helt enkelt beroende på hur man mäter, inte genom att på något vis förbättra mätapparaten. (Bland & Altman, 1986)

För att undersöka hur två mätmetoder de facto stämmer överens med varandra, med andra ord för att kolla ifall ett test kan bytas ut mot ett annat, utvecklade Martin Bland och Douglas Altman på 1980-talet ett enkelt statistiskt verktyg. Ursprungligen ville de ha ett verktyg för medicinska syften för att jämföra två mätmetoder som inte kan kalibreras med ett mätbart värde. Syftet var att se hur ett nytt test förhöll sig till ett gammalt och om det nya kunde användas istället för det etablerade. De drog slutsatsen att korrelation inte räckte och kunde leda till farliga medicinska följder (Bland & Altman, 1986) En direkt liknelse kan dras till prov och förhör som ett mått på elevers kunskap. Läraren kan inte granska elevens absoluta kunskapsmängd, än mindre hennes förmåga att tillämpa, verkställa och utnyttja den till nyskapande, så indirekta metoder måste användas.

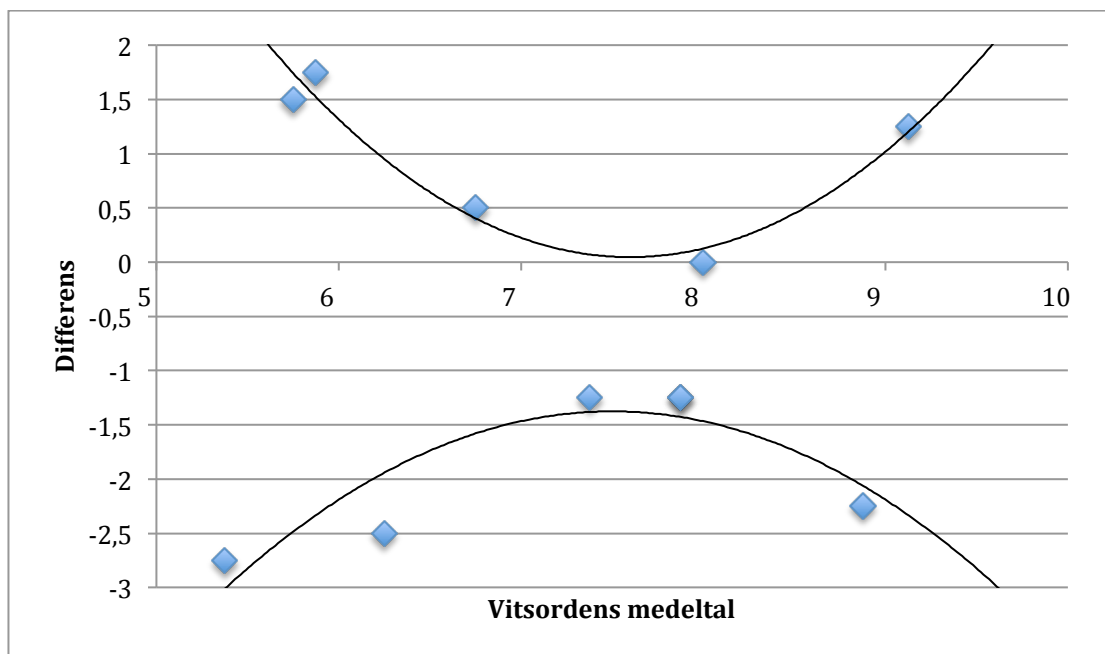
Då ett prov görs försöker man i praktiken mäta kunskap, men det finns ingen garanti för att det uppmätta resultatet stämmer överens med den eteriska kunskapen eleven besitter. Däremot är den provstruktur som oftast används i skolor och som användes för det här ändamålet, nämligen essäsvar, problemlösningar, definitioner och så vidare etablerade i vårt skolsystem. De kan anses utgöra guldstandarden för kunskap. För att införa ett nytt testsystem är det därför viktigt att undersöka huruvida det nya ger jämförbara resultat med det gamla och vad skillnaderna kan tänkas betyda.

Bland-Altman -metoden är i sig enkel; man jämför två uppsättningar data, A och B. Då blir x-axeln deras medeltal $[(A+B)/2]$ och y-axeln den absoluta skillnaden mellan värdena $(A-B)$. I det här fallet är A förhørsresultatet och B kartans resultat. (Dewitte, Fierens, Stöckl, & Thienpont, 2002) För en ordentlig jämförelse bör härtill räknas ut ett 95 % -säkerhetsområde $d \pm 2s$, där d är medeltalet av differensen och s är standardavvikelsen för differensen. (Bland & Altman, 1986). Eftersom prov inte kan tänkas vara riktigt så regelbundna kommer detta inte att göras i den här studien. Då

denna analys görs är grundskoleresultaten fortsättningsvis betydligt bättre än gymnasieresultaten, men man ser att också i grundskolan är det oroväckande många elever vars resultat ligger långt ifrån x-axeln, alltså vars resultat i de två olika testen ligger långt från varandra. Resultaten av en Bland-Altman analys av materialet presenteras i Graf 3 och 4.



Graf 3 Bland-Altman-graf för grundskolan. Den svarta kurvan utmärker en trend enligt vilken de verkligt svaga och de verkligt starka eleverna, alltså de med medeltal under 6 och över 8 i allmänhet gjorde bättre ifrån sig med begreppskartan medan eleverna med medeltalet mellan 6 och 9 i allmänhet klarade det traditionella förhåret bättre.



Graf 4 Bland-Altman-graf för gymnasiet. De svarta kurvorna markerar en trend i resultaten enligt vilken eleverna med ett medeltal på 6 ½ till 8 ½ verkar ligga mycket närmare varandra än de övriga.

6 Diskussion

Syftet med den här studien har varit att undersöka hur väl begreppskartor skulle kunna användas för att kompensera traditionella prov i grundskolan och i gymnasiet. Tanken är inte att alla prov skulle bytas ut utan att begreppskartor skulle användas som ett snabbt och effektivt medel för lärare att bedöma hur nytt material anammats av eleverna medan kursen pågår, delvis för att kunna täppa till brister i kunskapen, delvis för att skapa sig en bild av elevernas förståelse för stoffet för kursvitsordet. Att använda begreppskartor som bedömningsmetod skulle potentiellt ha den ökade fördelen att det skulle motivera elever att använda metoden som i upprepade empiriska studier har visat sig leda till bra resultat men som trots detta används av mycket få i våra skolor. För att få den fulla effekten av begreppskartornas fördelar borde lärarna naturligtvis också införa inläring, repetition och undervisning med hjälp av dem, men dessa aspekter faller utanför den här studien.

Fysikundervisningen i grundskolan och gymnasiet skiljer sig till sitt innehåll och sin karaktär kraftigt från varandra. Som det redan tidigare sagts har gymnasiet en benägenhet att vara inriktat på problemlösning i form av matematiska uppgifter och handlar till en stor del om att förstå och kunna tillämpa formler. I grundskolans fysikundervisning spelar matematiken en betydligt mindre roll vilket leder till att kurserna inte kan ha kvantitativ problemlösning som tyngdpunktsområde utan en mer kvalitativ förståelse av materialet. Det här syns mycket tydligt i hur de olika grupperna presterade i sina olika uppgifter och hur deras prestationer avspeglade sig i den större kursen. Begreppskartor är bäst på att mäta en holistisk förståelse för ett område, inte utantillinläring och inte kvantitativ problemlösning. Huruvida de ens avspeglar kvalitativ problemlösningsförmåga kan diskuteras. Därmed faller både gammaldags utbildning som till stor del handlade om att lära sig fakta utantill utanför begreppskartan och undervisning som endast handlar om problemlösning framom förståelse kommer att undervärdera de elever som förstår ämnet men har svårt att prestera matematiskt i det.

En överblick av begreppskartorna ger att gymnasisterna i många fall hade en betydligt mer hierarkisk, konkret, syn på materialet, vilket syns i att de flesta byggde kedjekartor med stor emphasis på länkar som handlade om formler eller att räkna ut saker. Det är inte en brist hos eleverna, det är vad gymnasiet kräver och visar snarast en brist i dagens gymnasietest där man genom att bolla med formler kan klara sig hyfsat bra i en kurs utan att ha större förståelse för de bakomliggande fenomenen. På grundskolesidan är förvisso samplet mer än 50 % större vilket i sig

formen av karta var även här kedjekartan men tätt därefter kom nätverket med många utförliga och välintegrerade nätverk. Något förvånande och eventuellt positivt var att navkartorna inte i någondera gruppen var särskilt vanliga. I allmänhet har navkartorna ansetts höra till de som nybörjare är mest benägna att konstruera (Kinchin et al., 2000) så det är ett bra tecken att så gott som alla elever i de två grupperna skapade åtminstone lite mer mångsidiga kopplingar, fast många var starkt centraliserade.

Vad beträffar begreppskartan som bedömningsmetod kan sägas att svårighetsmässigt verkar både det traditionella provet och begreppskartan hålla sig inom liknande marginaler med ett medeltal för det traditionella provet ett halvt till en fjärdedels vitsord högre än för begreppskartan. Fler, längre och större undersökningar skulle behövas för att fastställa de exakta förhållandena mellan de två formerna. Så att använda sig av begreppskartor skulle åtminstone inte enligt den här studiens resultat leda till stora kast i elevernas slutvitsord klassvis. Med andra ord kommer begreppskartan som bedömningsmetod att ge en liknande spridning mellan elevernas poängtal som de traditionella proven. De mäter något, men mäter de vad lärarna avser?

En korrelations- och överensstämmelseundersökning av de två gruppernas resultat visar tydliga skillnader. Korrelationen för grundskolegruppen är mycket hög medan den för gymnasiet är närmast obefintlig (se Graf 1 och 2). Mer erfarenhet av metoden och en fin slipning av bedömningskriterierna kunde tänkas öka korrelationen åtminstone i grundskolegruppens fall, eventuellt också i gymnasiet. Orsaken till att det är oklart huruvida det skulle hjälpa i gymnasiet är att det är osäkert om de traditionella proven överhuvudtaget mäter samma sak som begreppskartorna, så ingen form av slipning skulle hjälpa där. Den lilla korrelation man stöter på kan rätt enkelt förklaras med att starka elever har en tendens att vara starkare på alla akademiska uppgifter och har lättare att anamma och använda ett nytt system än de svagare, så en viss korrelation uppstår.

Bland & Altmans metod att jämföra ett nytt test med ett gammalt ger en liknande men upplysande bild av situationen (se graf 3). För grundskolan verkar det som om de verkligt svaga och de verkligt starka eleverna har en tendens att gynnas av begreppskartan medan eleverna med ett medeltal för uppgifterna på mellan 6 och 8 i allmänhet fick sämre resultat. En tolkning kunde vara att de svagaste eleverna gynnades av begreppskartan emedan det med den är lätt att få åtminstone lite poäng, medan de starkaste gynnades för att de hade en gedigen uppfattning av området. I bägge fallen skulle det på det stora hela vara positivt att använda begreppskartor som bedömningsmetod. De ger åt ena sidan de starkaste en chans att

uttömmande visa sin kunskap medan de ger de svagaste en möjlighet att prestera något. Den potentiellt negativa sidan är medeleverna som gjorde sämre (ibland betydligt sämre) ifrån sig i begreppskartsuppgiften. Förutom vissa aspekter som ovana vid provformatet kunde en förklaring vara att elever med vitsord runt 7 har en tendens att lära sig utantill. Det leder till bra provresultat emedan det är svårt att skapa ett prov som helt och hållet utesluter utantillinläring men deras bristande förmåga att egenhändigt greppa helheten blottläggs genom begreppskartan. Det negativa med traditionella prov är alltså att en stor del av eleverna kommer undan med endast utantillinläring. Det är en brist i bedömningskriterierna som väl kunde försöka åtgärdas. Här kunde utnyttjas begreppskartans förmåga att föra dessa brister i dagen, också för eleverna. Det kan i många fall vara så att de tror sig förstå när de egentligen bara kan.

På gymnasiesidan ger Bland & Altmans metod mycket mer splittrade resultat (se Graf 4) och fördelningen är svårare att tolka. Det verkar som om de två metoderna möts bäst för samma medelever, de med vitsord mellan 6 och 8, men också bland dessa är kastet enormt. Speciellt för gymnasiet måste det konstateras att med 10 datapunkter och en så pass stor spridning är det svårt att dra långtgående slutsatser, det skulle helt enkelt krävas ett större sampel. Den statistiska analysen talar dock sitt tydliga språk i frågan om begreppskartans relevans för den traditionella gymnasieundervisningen.

Som en ny instuderingsmetod är begreppskartan effektivast ju tidigare den presenteras för en individ och i gymnasiet kan många elever redan ha utvecklat ett inrotat sätt att behandla ny information som inte är lätt att tillämpa på begreppskartor. Grundskolans fokus på allmän förståelse för fenomen framom kvantitativ problemlösning hjälper ytterligare till att göra begreppskartan till ett mer passande bedömningsverktyg på de lägre nivåerna än de högre, åtminstone så som kurserna är uppbyggda idag.

Begreppskartan är ett empiriskt bekräftat effektivt sätt att lära sig och elever som kan utnyttja metoden klarar sig statistiskt sätt bättre i sina studier. Genom att förenkla bedömningen tillräckligt och ge lagom strama tyglar åt eleverna kan begreppskartan användas som förhörsmetod utan att det kräver mer tid än ett traditionellt förhör. Därtill ger begreppskartan läraren lättare och smidigare insikter i elevens tankeprocesser och kunskapsstrukturer. Den här studien har visat att i grundskolan kan en lärare mycket väl använda begreppskartan åtminstone som ett tillägg till de traditionella förhören utan att vara rädd för att gå miste om något väsentligt eller slösa bort värdefull tid. Likaså har studien visat att det samma inte är fallet i gymnasiet och att en lärare på den nivån inte kan vänta sig resultat som är

direkt relevanta för kurserna, åtminstone i frågan om hur de traditionellt bedömts. Gymnasiefysiken verkar allt för inriktat på formeljonglering för att begreppskartan skall vara till stor hjälp. Huruvida begreppskartor trots allt kunde ha en plats i gymnasiet som inlärnings- och repetitionsmetod faller utanför den här studiens ramar men det kunde tänka sig att en elev som lärt sig tekniken i grundskolan väl kunde tillämpa den också i gymnasiet.

Källförteckning

- Acton, W., Johnson, P., & Goldsmith, T. (1994). Structural Knowledge Assessment. *Journal of Educational Psychology*, 86(2), 303–311.
- Bland, M., & Altman, D. (1986). statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 307–310.
- Derbentseva, N., Safayeni, F., & Cañas, A. J. (2007). Concept maps: Experiments on dynamic thinking. *Journal of Research in Science Teaching*, 44(3), 448–465. doi:10.1002/tea.20153
- Dewitte, K., Fierens, C., Stöckl, D., & Thienpont, L. (2002). Application of the Bland–Altman Plot for Interpretation of MethodComparison Studies. *Clinical Chemistry*, 48(5), 799–801.
- Feynman, R. & Gleick, J. (1994). The Character of Physical Law, Introduction by James Gleick. *Modern Library, New York*.
- İngeç, Ş. K. (2009). Analysing Concept Maps as an Assessment Tool in Teaching Physics and Comparison with the Achievement Tests. *International Journal of Science Education*, 31(14), 1897–1915. doi:10.1080/09500690802275820
- Kinchin, I. M., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research*, 42(1), 43–57. doi:10.1080/001318800363908
- McClure, J., Sonak, B., & Suen, H. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36, 475–492.
- Nesbit, J., & Adesope, O. (2006). Learning With Concept and Knowledge Maps. *Review of Educational Research*, (Fall 2006, Vol. 76), 413–448.
- Nicoll, G. (2001). A three-tier system for assessing concept map links: a methodological study. *International Journal of Science Education*, 23(8), 863–875. doi:10.1080/09500690010025003
- Novak, J. & Gowin, B. (1984). Learning how to Learn. *Cambridge University Press*.
- Ruiz-Primo, M. (2000). On the Use Of Concept Maps As An Assessment Tool in Science:. *Revista Electrónica de Investigación Educativa*, 2(1), 29–53.
- Ruiz-Primo, M., & Shavelson, R. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569–600.
- Ruiz-Primo, M., & Shavelson, R. J. (1997). Concept-map based assessment, 1–52.
- Ruiz-Primo, M., Schultz, S., Li, M., & Shavelson, R. (2001). Comparison of the Reliability and Validity of Scores from Two Concept-Mapping Techniques. *Journal of Research in Science Teaching*, 260–278.
- Safayeni, F., Derbentseva, N., & Cañas, A. J. (2005). A theoretical note on concepts and the need for Cyclic Concept Maps. *Journal of Research in Science Teaching*, 42(7), 741–766. doi:10.1002/tea.20074
- Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching*, 42(2), 166–184. doi:10.1002/tea.20049

Bilaga 1: Lista på begrepp samt förhöret för grundskolegruppen

Prov i vågrörelse och ljud

/24

Tid: 45 minuter

Namn:

Efter att du lämnat in får du nästa uppgift, att göra en begreppskarta. Begreppskartan är en helt skild uppgift, så att du förklarat något i den här delen av provet påverkar inte begreppskartan.

1. Förklara skillnaderna eller likheterna.
 - a. Vad är skillnaden mellan en transversell och en longitudinell vågrörelse?
 - b. Vad har vår hörsel gemensamt med knuffandet av en kompis på en gunga?
 - c. Vad är skillnaden mellan frekvens och period?
2. Förklara följande fenomen.
 - a. Vilken egenskap hos pendeln (som Galileo Galilei upptäckte) gör att den brukar användas i klockor?
 - b. Då du metar skvalpar vågorna ständigt mot stranden, men metspöets flöte flyter inte genast in mot land. Varför?
 - c. En hundvissla är en visselpipa som hundar hör men inte vi och som kan användas för att straffa olydiga hundar. Enligt vilken princip fungerar hundvisslan? Kunde en liknande apparat skapas för skolbruk? Varför?
 - d. Under vissa omständigheter rör sig ljudet snabbare än i andra, förklara vilka dessa omständigheter är och ge en förklaring på varför detta händer. Rita bild om det hjälper.
3. Förklara följande ord:
 - a. Oscillator
 - b. Amplitud
 - c. Medium
 - d. Resonans
4. Räkna följande. Kom ihåg att $f = \frac{1}{T}$
 - a. Vad är frekvensen ifall perioden är 0,1 sekunder?
 - b. Hur lång period har ett system vars frekvens är 50 Hz.
 - c. Hjärtat på en mus slår 500 gånger i minuten. Vad är mushjärtats frekvens?

Begreppskarta

Konstruera en begreppskarta med de 10 givna begreppen.

- Inga andra begrepp får användas
- Koppla ihop begreppen med pilar med etiketter, länkar utan pil **och** etikett kommer inte att räknas!

1. Amplitud
2. Frekvens
3. Medium
4. Oscillator
5. Period
6. Resonans
7. Svängningsrörelse
8. Volym
9. Våglängd
10. Vågrörelse

Försök använda alla begreppen!

Poängsättning:

Felaktigt påstående	0 poäng
Korrekt men irrelevant påstående	1 poäng
Bristfälligt korrekt påstående	2 poäng
Korrekt påstående	3 poäng
Uttömmande korrekt påstående	4 poäng

Bilaga 2: Lista på begrepp samt förhöret för gymnasiegruppen

Förhör i Fysik 6 (Ellära)

namn: _____

A. Gör en mind map / tankekarta utgående från följande elva begrepp:

1. Elström
2. Resistans
3. Inre resistans
4. Källspänning
5. Polspänning
6. Effekt
7. Resistivitet
8. Batteri / spänningskälla
9. Elektrisk effekt
10. Motstånd
11. Joules lag

B. Förklara (kort) följande begrepp:

1. Ekvipotentialeyta
2. Oscilloscop
3. Kortslutningsström
4. En JORDAD krets
5. Diod

Bilaga 3: Provresultat

Elev	Kartan	Provet
A	6	7+
B	6	7 ½
C	10+	9+
D	6 ½	7
E	5+	5
F	7-	8+
G	10+	9 ½
H	10+	9-
I	5+	4
J	8 ½	8+
K	6+	6+
L	7	9
M	8	10-
N	4	4
O	6-	5
P	4	5 ½

Tabell 1 Grundskoleresultat

Elev	Kartan	Provet
A	7+	8 ½
B	5	7 ½
C	4	7-
D	8-	10
E	7-	5
F	7+	8 ½
G	7	6 ½
H	10-	8 ½
K	7-	8
L	6 ½	5
M	8	8

Tabell 2 Gymnasieresultat